

The actinome of *Dictyostelium* amoebae: comparative *in silico* and *in vivo* characterization

Jayabalan M Joseph

2014

LMU



Aus der Anatomischen Anstalt
Lehrstuhl: Anatomie III - Zellbiologie
Ludwig-Maximilians-Universität München
Vorstand: Univ. Prof. Dr. Michael Kiebler

Thema der Dissertation

**The actinome of *Dictyostelium* amoebae:
comparative *in silico* and *in vivo* characterization**

Dissertation
zum Erwerb des Doktorgrades der Humanbiologie
an der Medizinischen Fakultät der
Ludwig-Maximilians-Universität München

vorgelegt von
Jayabalan M Joseph
aus
Srivilliputtur, Indien

2014

Mit Genehmigung der Medizinischen Fakultät der Universität München

Berichterstatter: Prof. Dr. Michael Schleicher

Mitberichterstatter: Prof. Dr. Axel Imhof
Priv. Doz. Dr. Wolfgang Fischer
Priv. Doz. Dr. Anton Eberharder

Dekan: Prof. Dr. med. Dr. h.c. M. Reiser, FACR, FRCR

Tag der mündlichen Prüfung: 17.10.2014

Eidesstattliche Versicherung

Joseph M. Jayabalan

Name, Vorname

Ich erkläre hiermit an Eides statt,

dass ich die vorliegende Dissertation mit dem Thema

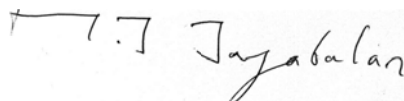
The actinome of *Dictyostellium amoebae*: comparative *in silico* and *in vivo* characterization.

selbständig verfasst, mich außer der angegebenen keiner weiteren Hilfsmittel bedient und alle Erkenntnisse, die aus dem Schrifttum ganz oder annähernd übernommen sind, als solche kenntlich gemacht und nach ihrer Herkunft unter Bezeichnung der Fundstelle einzeln nachgewiesen habe.

Ich erkläre des Weiteren, dass die hier vorgelegte Dissertation nicht in gleicher oder in ähnlicher Form bei einer anderen Stelle zur Erlangung eines akademischen Grades eingereicht wurde.

Nagercoil, 17.04.2013

Ort, Datum



Unterschrift Doktorandin/Doktorand

Table of contents:

List of publications.....	9
Summary.....	10
Zusammenfassung.....	11
1 Introduction.....	12
1.1 The cytoskeleton in eukaryotic cells.....	12
1.1.1 Microfilaments.....	12
1.1.2 Microtubules.....	13
1.1.3 Intermediate filaments.....	14
1.2 Actin binding proteins (ABP).....	15
1.3 Motor proteins.....	17
1.4 Prokaryotic cytoskeleton.....	17
1.5 Actin.....	18
1.5.1 Mutant actin.....	19
1.6 Actin related proteins (Arps).....	20
1.7 <i>D. discoideum</i> is an excellent model system to study the cytoskeleton.....	20
1.8 Aims of the project.....	21
2 Materials and methods.....	22
2.1 Materials.....	22
2.1.1 Software packages.....	22
2.1.2 Servers.....	23
2.1.3 Reagents.....	24
2.1.4 Centrifuges and rotors.....	25
2.1.5 Media.....	25
2.1.6 Buffers.....	27
2.1.7 Instruments.....	28
2.1.8 Microbial strains.....	29
2.1.9 Vectors.....	29

2.2 Molecular Biological Methods.....	30
2.2.1 Isolation of <i>D. discoideum</i> genomic DNA using a high pure PCR template preparation kit (Roche).....	30
2.2.2 Polymerase chain reaction (PCR).....	30
2.2.3 Agarose gel electrophoresis.....	31
2.2.4 Determination of DNA concentration.....	32
2.2.5 Isolation of DNA fragments from agarose gels.....	32
2.2.6 Purification of PCR products.....	32
2.2.7 DNA cleavage with restriction enzymes.....	33
2.2.8 Mini preparation of plasmid DNA using Qiagen miniprep kit.....	33
2.2.9 Mini preparation of plasmid DNA by the method of Holmes and Quigley.....	34
2.2.10 Phosphatase treatment.....	34
2.2.11 Ligation of DNA into a plasmid vector.....	34
2.2.12 Preparation of chemically competent cells.....	34
2.2.13 Transformation of <i>E. coli</i>	35
2.2.14 <i>E. coli</i> permanent cultures.....	35
2.2.15 Maxi preparation of plasmid DNA.....	35
2.3 Biochemical methods.....	36
2.3.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE).....	36
2.3.2 Commassie blue staining of proteins.....	37
2.3.3 Western blotting.....	37
2.3.4 Expression of GST-tagged protein.....	38
2.4 Cell biological methods.....	39
2.4.1 Growth in liquid medium and on agar plates.....	39
2.4.2 Preservation of spores.....	39
2.4.3 Storage of <i>D. discoideum</i> cells at -70°C.....	40
2.4.4 Transformation of <i>D. discoideum</i> cells.....	40
2.4.5 Cloning of transformants.....	41

2.4.6 Affinity purification.....	41
2.4.7 Indirect immunofluorescence.....	41
3 Results.....	43
3.1 The <i>D. discoideum</i> actinome.....	43
3.2 Genomic locations of selected actin genes.....	45
3.3 Phylogeny of the <i>D. discoideum</i> actinome.....	47
3.4 Actin and arps structural features, models, and sequence logos.....	54
3.4.1 Domains.....	56
3.4.2 Motifs.....	56
3.5 Localization signals and sorting motifs.....	61
3.5.1 Nuclear localization signal (NLS).....	61
3.5.2 Nuclear export signal (NES).....	61
3.5.3 Signal peptide prediction (SignalP).....	64
3.5.4 Transmembrane hidden markov model (TMHMM).....	64
3.5.5 Proline(P) glutamic acid(G) serine(S) threonine(T) rich regions (PEST motifs).....	64
3.6 Actin homologues across species.....	65
3.7 Actin related proteins (Arps).....	68
3.8 <i>D. discoideum</i> actin in comparison to actins from other organisms.....	70
3.9 Actin mutations that lead to human diseases, in relation to distinct actin variations in the <i>D. discoideum</i> actinome.....	71
3.10 Actin related protein 5 (Arp5).....	73
3.11 Actin related protein 6 (Arp6).....	82
4 Discussion.....	88
4.1 Actin and protein evolution.....	88
4.1.1 Actinome.....	88
4.1.2 Actin phylogeny.....	89
4.1.3 Actin related proteins.....	89

4.1.4 Structural evolution of actin.....	90
4.1.5 Convergent evolution of actin.....	90
4.1.6 Actin's structural motifs.....	90
4.1.7 Signals and motifs.....	91
4.1.8 Arp5 and Arp6.....	91
4.1.9 Disease mutants of actin.....	92
4.1.10 Post-translational modifications in <i>D. discoideum</i> actins.....	93
4.1.11 Unique characters of actin.....	93
4.2 <i>D. discoideum</i> and cellular evolution.....	95
4.2.1 Motility.....	95
4.2.2 Multicellularity.....	96
4.2.3 Cell differentiation.....	96
4.2.4 Cheaters.....	97
4.2.5 Tumor suppressors.....	97
4.2.6 Cell-cell adhesion.....	97
4.2.7 Phagocytosis.....	98
4.3 A sketch on the evolution of organisms.....	98
4.3.1 History of evolution.....	98
4.3.1.1 Neo-darwinism.....	98
4.3.1.2 Punctuations in fossils.....	99
4.3.1.3 Neo-darwinian speciation.....	99
4.3.1.4 Morphological discontinuity.....	99
4.3.1.5 Artifacts theory.....	99
4.3.1.6 Punctuations in taxa.....	100
4.3.1.7 Alternatives to neo-darwinism.....	100
4.3.2 Concepts in evolution.....	100
4.3.2.1 Micro-mutation and micro-evolution.....	100
4.3.2.2 Macro-mutation and macro-evolution.....	100

4.3.2.3 Fast and slow evolution.....	101
4.3.2.4 Environments role.....	101
4.3.2.5 Species branding factors & (Hypothesis of species branding).....	102
4.3.2.6 Quantitative traits.....	103
4.3.2.7 Species concepts.....	104
4.3.2.8 Prospects of micro-mutation.....	106
4.3.2.9 Prospects of macro-mutation.....	106
4.3.2.10 Evolution in action.....	107
4.3.3 Cambrian explosion.....	108
4.3.3.1 Mega evolution / super evolution / enhanced evolution.....	109
4.3.3.2 Evo-nhancer hypothesis.....	110
4.3.3.3 Support for mega-evolution.....	111
4.3.3.4 Advantages of evo-nhancer hypothesis.....	112
4.3.4 Resurfacing of past theories.....	113
4.3.4.1 Neo-lamarckism and epigenic inheritance.....	113
4.4 Natural history of earth.....	114
5 References.....	117
List of figures.....	133
List of tables.....	134
Abbreviations.....	135
Acknowledgements.....	139

List of publications

Original publications:

Jayabalan M. Joseph, Petra Fey, Nagendran Ramalingam, Xiao I. Liu, Meino Rohlf, Angelika A. Noegel, Annette Müller-Taubenberger, Gernot Glöckner, Michael Schleicher. The Actinome of *Dictyostelium discoideum* in comparison to actins and actin-related proteins from other organisms. PLOS one – 09 July 2008 | Volume 3 | Issue 7 | e2654

Rajesh Arasada, Annika Gloss, Budi Tunggal, **Jayabalan M. Joseph**, Daniela Rieger, Subhanjan Mondal, Michael Schleicher and Angelika A. Noegel. Profilin isoforms in *Dictyostelium discoideum*. Biochimica et Biophysica Acta (BBA) - Molecular Cell Research May 2007, 631-641

Poster presentations:

Jayabalan M. Joseph, Petra Fey, Nagendran Ramalingam, Xiao I. Liu, Meino Rohlf, Angelika A. Noegel, Annette Müller-Taubenberger, Gernot Glöckner, Michael Schleicher. The actinome of *Dictyostelium discoideum* in comparison to actins and actin-related proteins from other organisms. 2008, CSH Meeting 'Biology of Genomes', Cold Spring Harbor Laboratories, U.S.A.

Jayabalan M. Joseph, Petra Fey, Nagendran Ramalingam, Xiao I. Liu, Meino Rohlf, Angelika A. Noegel, Annette Müller-Taubenberger, Gernot Glöckner, Michael Schleicher. The *D. discoideum* actinome. 2007, International Dictyostelium Meeting, Rotenburg/Fulda, Germany.

Annette Müller-Taubenberger, **Jayabalan M. Joseph**, Marlis Fürbringer, Angelika A. Noegel, Michael Schleicher. Actin-related proteins in *Dictyostelium discoideum*. 2007, Symposium CRC Chromatin Assembly and Inheritance, Adolf Butenandt Institute, Munich, Germany.

Summary

The highly conserved protein actin is the building block in the cytoskeleton of eukaryotic cells and provides a structural framework known as the microfilament system. The molecular principle of actin-based amoeboid movement was so successful in evolution that it was kept nearly identically from lower (e.g. amoebae) to higher (e.g. neutrophils) eukaryotes. To understand this type of cellular movement one has first to identify and to characterize the proteins which play a major role during the dynamic rearrangement of actin. The collection of actin isoforms, of actin-variants and actin related proteins (Arps) in a given cell is known as the 'actinome' whose number of proteins can be quite different from one organism to the next. Therefore, the present work describes studies on the actinome of the social amoeba *Dictyostelium discoideum*, compares the findings with actinomes from other organisms, and discusses similarities and alterations that might have happened during evolution.

D. discoideum is among the oldest organisms which exhibit actin-based amoeboid movement, the genome is completely sequenced and the system can be easily studied by molecular and biochemical approaches. The study was started using bioinformatics and the computational methods provided a global view on the *D. discoideum* actinome. It turned out that the *D. discoideum* genome comprises a total of 33 actin and 8 Arp genes, seven actin genes are putative pseudogenes. Interestingly, there are 17 distinguishable actin genes which code for identical proteins. Phylogenetic analyses helped to understand the putative duplication events during evolution. Modelling of the three-dimensional structures showed that the typical actin-fold, the ATP-binding pocket, and other functional domains are highly conserved. Homologues of the members of the *D. discoideum* actinome across various model organisms clearly demonstrated which amino acids in conserved domains are of special importance. All Arp subfamilies that are found in mammals are also present in *D. discoideum*. Two of the actin related proteins, Arp5 and Arp6, were selected for molecular and cellular studies. Using fluorescently labeled fusion proteins first data indicated that both Arps are present also in the nucleus, suggesting an involvement in chromatin reorganization.

Zusammenfassung

Das hochkonservierte Protein Aktin ist die grundlegende Einheit des Zytoskeletts in eukaryontischen Zellen und formt das als Mikrofilamentsystem bezeichnete intrazelluläre Gerüst. Das molekulare Prinzip der amoeboiden Fortbewegung basiert auf Aktin und war so erfolgreich in Lauf der Evolution, dass es nahezu unverändert geblieben ist von niederen (z.B. Amöben) zu höheren (z.B. Neutrophilen) Eukaryonten. Um diesen Typ der zellulären Fortbewegung verstehen zu können, muss man zuerst die Proteine identifizieren und charakterisieren, die eine Hauptrolle bei den dynamischen Umbauvorgängen im Aktin Zytoskelett spielen. Alle Aktine, Aktin-Isoformen und Aktin-verwandten Proteine ('actin related proteins', 'Arps') werden als 'Aktinom' zusammengefasst, dessen Proteinspektrum sich in den verschiedenen Organismen stark unterscheiden kann. Die vorliegende Arbeit sollte deshalb das Aktinom der sozialen Amöbe *Dictyostelium discoideum* charakterisieren, die Ergebnisse mit den Aktinomen anderer Organismen vergleichen, und Ähnlichkeiten bzw. Unterschiede, die während der Evolution aufgetreten sind, diskutieren.

D. discoideum gehört zu den ältesten Organismen, deren amoeboide Fortbewegung auf Aktin beruht, das Genom ist vollständig sequenziert, und das Modell-System kann für molekulare und biochemische Versuchsansätze verwendet werden. Bioinformatik und Rechner-basierte Methoden führten zu einem weitgehenden Einblick in das *D. discoideum* Aktinom. Es zeigte sich, dass das *D. discoideum* Genom 33 Aktin- und 8 Arp-Gene enthält, 7 Aktin-Gene wurden als vermutliche Pseudogene eingestuft. Interessanterweise wurden 17 unterscheidbare Aktin-Gene gefunden, die für identische Proteine kodierten. Phylogenetische Analysen haben geholfen, die vermutlichen Genduplikationen im Lauf der Evolution zu verstehen. Die Modellierung der dreidimensionalen Strukturen gab wichtige Einblicke, inwieweit die typische Aktinstruktur, die ATP-Bindungs tasche, und andere funktionelle Domänen erhalten geblieben waren. Homologe der Proteine aus dem *D. discoideum* Aktinom in anderen Modellorganismen demonstrierten ganz klar, welche Aminosäuren in konservierten Domänen von besonderer Bedeutung sind. Darüber hinaus konnte festgestellt werden, dass alle Arp Unterfamilien in *D. discoideum* auch in Säugern vorhanden sind. Zwei der Aktin-verwandten Proteine, nämlich Arp5 und Arp6, wurden für weitergehende molekulare und zelluläre Untersuchungen ausgewählt. Mit Hilfe von GFP-markierten Fusionsproteinen konnte gezeigt werden, dass beide Arps auch im Nukleus vorhanden und dort wohl an der Umorganisation des Chromatins beteiligt sind.

1. Introduction

1.1 The cytoskeleton in eukaryotic cells

The cytoskeleton is the structural frame work of a cell. It is a dynamic structure which gives the cell its shape and enables cellular motion. It also acts as a highway for intracellular transport. The typical eukaryotic cytoskeleton is usually composed of microfilaments, microtubules and intermediate filaments. All cytoskeletons contain in addition a number of associated proteins that either support the overall structure or regulate cytoskeletal dynamics(Kreis and Vale, 1999). Furthermore, the individual cyoskeletons are not separated and entirely independent structures but they cooperate in many fundamental cellular activities. A classical example is a cell division which requires the interplay between the microtubular system for the separation of the chromosomes and the actin cytoskeleton that organizes finally the cleavage furrow to guarantee the separation of the emerging daughter cells (Fig. 1).The following descriptions of the three major cytoskeletal systems can only summarize the most characteristic features.

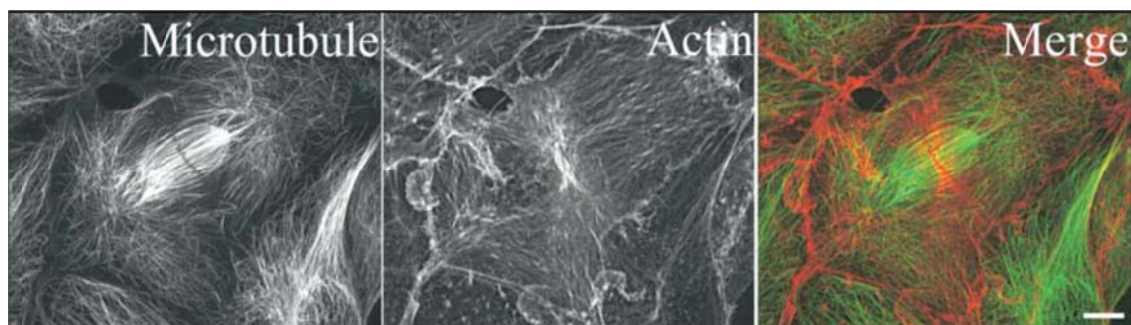


Figure 1: Organization of the microtubular and the actin system during cytokinesis. Using a porcine kidney proximal tubule cell line (LLC-PK1) microtubules were stained with an anti-tubulin antibody (green) and microfilaments with rhodamine-phalloidin (red). The merge clearly shows the location of F-actin in the periphery and the emerging cleavage furrow, whereas the microtubules accumulate especially in the spindle (Source: (Murthy and Wadsworth, 2005), scale bar = 10 μ m).

1.1.1 Microfilaments

Microfilaments are about 5-9nm in diameter and composed of two strands of actin chains. The filaments form a mesh lining at the inner periphery of a cell, thereby resisting tension and maintaining cell shape. They also organize distinct structures on

the cell membrane like filopodia, lamellipodia and invaginations, and are also known to participate in cell to cell and cell to matrix junctions (Popp and Robinson, 2011).

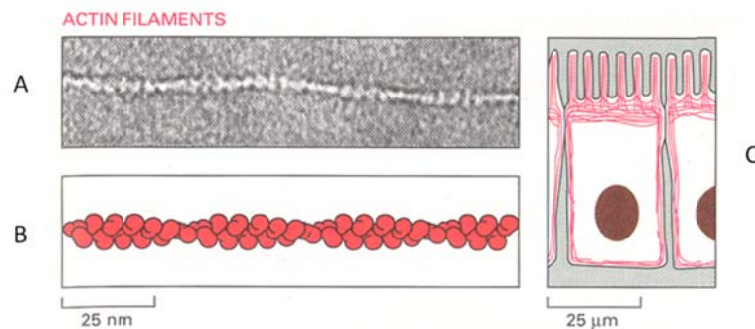


Figure 2: Actin filaments. (A) Electron micrograph of an actin filament. (B) Schematic representation of an actin filament as a two-start helix, and (C) of an epithelial cell showing the location of actin filaments at the periphery (Source: (Alberts, 1994)).

1.1.2 Microtubules

Microtubules are cylindrical structures of about 24 nm in diameter and built from 13 linear protofilaments, which in turn are made of alpha/beta tubulin subunits. With respect to polymerization, microtubules are polymerized with the help of GTP while microfilaments require ATP (Lloyd, 2011). Microtubules start from the centrosome (minus ends) and radiate towards the cell periphery (plus ends). They act as tracks for motor proteins which direct organelle transport. Triplet sets form centrioles and basal bodies which are the sites of nucleation for microtubules. Nine doublets organized around two microtubules give rise to cilia and eukaryotic flagellae which play their role in cell motility. Microtubules play also a vital role during cell division as the mitotic spindle is made of microtubules. They also help in the synthesis of cell walls in plants (Paredes *et al.*, 2006).

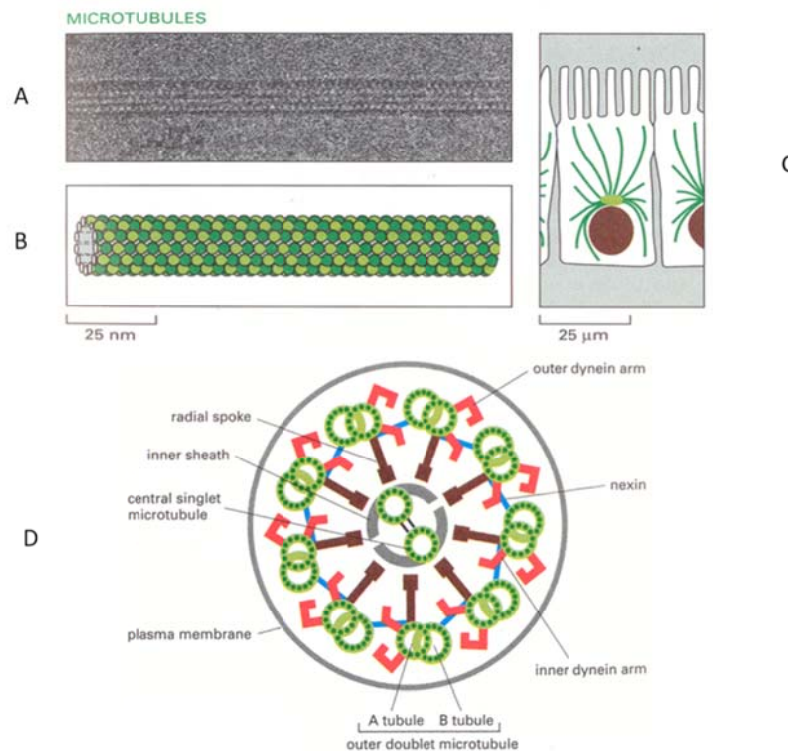


Figure 3: Microtubules. (A) Electron micrograph of a microtubule. (B) Schematic drawing of a microtubule, (C) schematic figure of cells with microtubules radiating from the centrosome. (D) Scheme of a cross section of a flagella showing a 9x2+2 arrangement of microtubules (Source: (Alberts, 1994)).

1.1.3 Intermediate filaments

Intermediate filaments are about 10 nm in diameter and got their names as their sizes are in between the diameters of microfilaments and microtubules. Unlike the former ones which are composed of one specific protein, intermediate filaments are made up of many different, often tissue-specific types of proteins. Examples of intermediate proteins are nuclear lamins, vimentin like proteins, keratins and neuronal intermediate filaments (Herrmann and Strelkov, 2011).

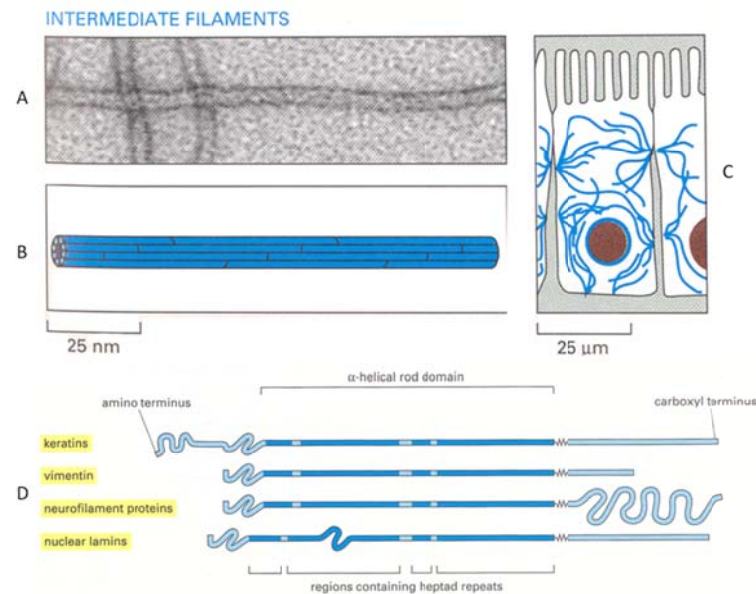


Figure 4: Intermediate filaments. (A) Electron micrograph of intermediate filaments. (B) Schematic drawing of an intermediate filament. (C) Schematic figure of a cell with intermediate filaments which are in contact with supramolecular structures at the cell membrane. (D) Domain structure of various intermediate filaments (Source: (Alberts, 1994)).

1.2 Actin binding proteins (ABP)

Actin is known to recruit a maximum number of binding partners (Dominguez, 2004). Most of the ABPs obviously bind to the exposed hydrophobic cleft of actin and thus affect the conformation of the actin molecule. As summarized in Fig. 5, actin-binding proteins cover a huge number of functions that are often tightly regulated by upstream signaling cascades.

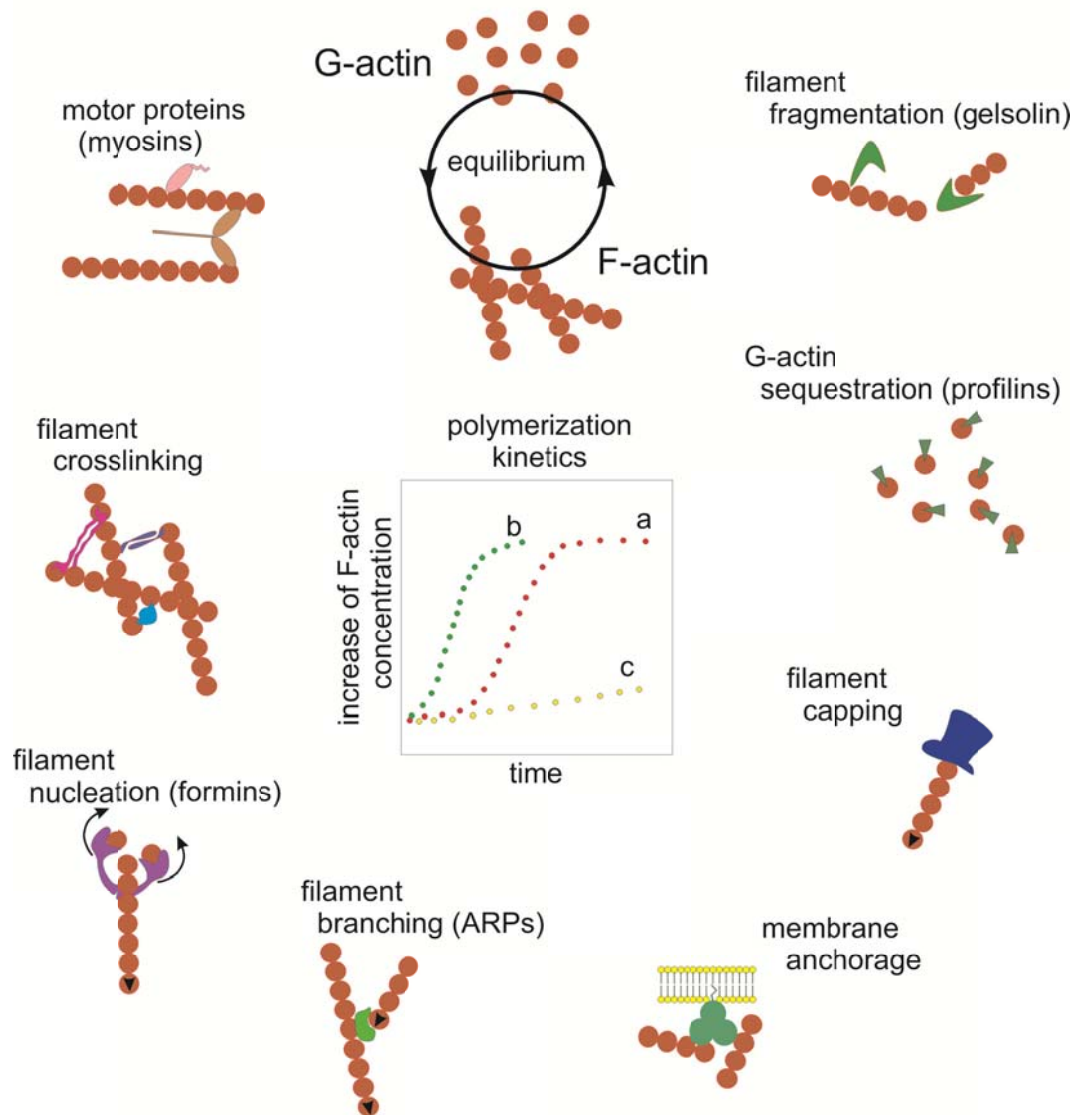


Figure 5: The impact of actin binding proteins on polymerization kinetics and stability of the microfilament network. The essential feature of actin dynamics is the nearly 50:50 equilibrium of monomeric (G-actin) and filamentous (F-actin) in a living cell. Normal polymerization kinetics (a, middle panel) can be stimulated by nucleation (b) if the cell needs instantaneously more actin filaments, or inhibited (c) if low viscosity is required in the cytoplasm. Several families of actin binding proteins directly influence these kinetics. Severing proteins reduce the number of filaments, sequestering proteins lower the concentration of polymerizable actin, capping proteins inhibit filament elongation, or formins enhance the appearance of actin filaments by nucleation. Other actin binding proteins act as motor proteins, or stabilize existing networks by filament branching and crosslinking, or association with the membrane (Rajesh, 2006).

1.3 Motor proteins

Motor proteins are a group of proteins that translate the chemical energy in the cell into movement. This could result in intracellular movements like transport of organelles and vesicles, cytokinesis or other motile activities like those of ciliated cells and muscles. Motor proteins are classified on the type of cytoskeletal structure they travel on. Myosin interacts with microfilaments; dynein and kinesin are motors that interact with microtubules. Motors that interact with intermediate filaments are not yet known. The typical structure is a head region which is the domain that binds to the filament, neck and a tail which vary in length and sequence. Molecular structures are known for myosin and kinesin which superimpose over the head region but differ in their amino acid sequence. A common ancestor is believed to have given rise to at least 12 classes of myosins which are further diversified by different tail topologies. Motor proteins are directional in motion. Myosin moves towards the barbed end of microfilaments. Kinesin usually moves towards the plus end of microtubules while dynein moves towards the minus end. There are also several groups of motor associated proteins (Kreis and Vale, 1999).

1.4 Prokaryotic cytoskeleton

The prokaryotic cytoskeletal proteins though they are not similar to their eukaryotic counterparts in protein sequence, share structural similarities and could have given rise to their eukaryotic counterparts through convergent routes (Shih and Rothfield, 2006). Structural homologs for actin, tubulin and intermediate filaments have been found; there is also a MinD-ParA (Lutkenhaus, 2012 Sep) group which is unique to bacteria. The actin homologs differ in sequence to the eukaryotic counterparts but share the actin fold. MreB, ParM and MamK (Carballido-Lopez, 2006) are the typical actin homologs. MreB also known as murein cluster B is the second gene of the Mre operon (Kruse *et al.*, 2005). This operon is involved in the synthesis of murein which forms the peptidoglycan layer thereby giving the cell its shape. Conditional mutants are available that make the rod shaped bacteria turn round. MreB polymerizes equally well in the presence of ATP and GTP. MreB also plays a role in the polarity of the cell by determining the location of bacterial flagellae. They are also involved in chromosomal segregation. ParM (van den Ent *et al.*, 2002; Moller-Jensen *et al.*, 2003) is another actin homolog which together with ParR is involved in plasmid partitioning by

physically separating the plasmids in dividing cells. MamK(Komeili *et al.*, 2006; Taoka *et al.*, 2007) is an actin homolog found in *Magnetospirillum magneticum* which is involved in the subcellular organization of the membranes of magnetosomes. FtsZ and BtubA/B(Schlieper *et al.*, 2005) are tubulin homologs. FtsZ(Mingorance *et al.*, Aug. 2011) forms a Z-ring(Stricker and Erickson, 2003; Srinivasan *et al.*, 2008) that enables cell division. BtubA / BtubB(Takeda *et al.*, 2008; Sontag *et al.*, 2009) are found only in the genus *Prostheco bacter*. Just like eukaryotic tubulin they polymerize in the presence of GTP. Crescentin(Esue *et al.*, Jan. 2010) is so far the only intermediate filament protein homolog found in prokaryotes which shows up to 40% similarity to eukaryotic IFs. The CreS(Margolin, 2004) protein gives the typical comma shape to *C. crescentus* which, when mutated, results in the loss of the curved shape.

MinD(Shen and Lutkenhaus, Nov. 2010) is involved in septal placement in *E. coli*, together with minC and minE genes of the min operon. ParA and ParB(Bignell and Thomas, 2001; Havey *et al.*, 2012 Jan) proteins are involved in plasmid partitioning. There is also a group of proteins that are recently found, which can form extended helical polymers. They are SetB, Sec proteins(Shiomi *et al.*, 2006) and Tar.

1.5 Actin

Actin and its associated proteins are major constituents of the cell, with around 25 % in non-muscle cells and 60% in muscle cells(Kreis and Vale, 1999). Being the structural basis of a cell, actin is ubiquitous and highly conserved like the histones, which form the structural basis of the chromatin(Alberts, 1994). Actin exists as a monomer called G-actin but can polymerize to filaments called F-actin, which takes place by the interplay of many actin binding proteins, Mg^{2+} and ATP (see also Fig. 5). The actively growing ends of the filament are called plus ends (barbed ends) while the lagging end is the minus end (pointed end), each end has its own array of ABPs which play their role in the actin dynamics. In the actin structure model (Kabsch and Holmes, 1995) the subdomains 1 and 2 of the molecule are believed to be exposed on the outer surface of the actin filament and the 3 and 4 oriented towards the core of the filament. Especially biochemical approaches and genome sequencing have shown that conventional actin is present in all eukaryotes and absent in eubacteria and archaea, with the sequences varying slightly across the species. Actin shows some structural

similarity to HSP70 (Flaherty *et al.*, 1991) and hexokinase which suggests their common origins. Warm blooded animals have six classical isoforms of actin which can be classified by isoelectric focusing as α -, β - and gamma γ -actins. Added to this they are further distinguished based on their location as α -skeletal muscle, α -aortic smooth muscle, α -cardiac muscle, γ 2-enteric smooth muscle and cytoplasmic isoforms. Myoactin (Alberts, 1994) is found in sarcomeres which forms the functional core of skeletal muscles. An ordered array of actin filaments and myosin bundles together with other proteins like titin provide the muscle power in the presence of ATP and Ca^{2+} . Actin monomer has four structural domains and five functional motifs (Dayel and Mullins, 2004). There are two phosphate binding motifs, two connecting motifs and one adenosine binding motif.

1.5.1 Mutant actin

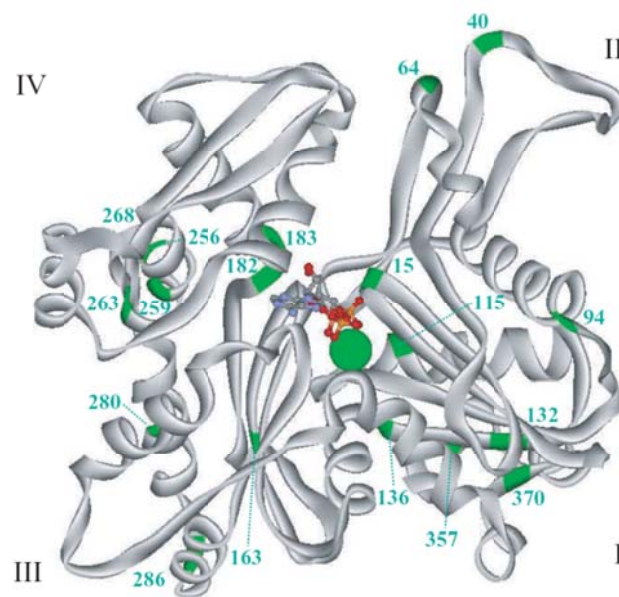


Figure 6: Relatively frequent mutations in the actin molecule: ActA1 mutation sites are marked on the structure of wild type actin (Costa *et al.*, 2004).

Many mutant actins have been reported in yeast and *Drosophila* and several in mammalian cell lines (Kreis and Vale, 1999). Naturally occurring ActA mutations in humans have been reported, which range from mild to severe muscle defects. The severity depends upon the site of mutation on the ActA gene (Costa *et al.*, 2004). ActB

mutations are known to cause severe developmental defects (Procaccio *et al.*, 2006). ActG mutations are reported in resulting in cardiac hypertrophy (Olson *et al.*, 1998). An oncogenic actin mutation has also been reported (Kakunaga *et al.*, 1984). Tools for expression of mutant actins as fusion proteins have been developed in *D. discoideum* (Noguchi *et al.*, 2007).

1.6 Actin related proteins (Arps)

Arps are a group of proteins that show varying degrees of similarity with actin and a few of them have preserved the actin structural folds. They have been reported from algae to humans but the presence and copy numbers of Arp subtypes vary (Muller *et al.*, 2005; Joseph *et al.*, 2008). Actins and Arps are believed to have risen from a common ancestor (Kreis and Vale, 1999). Arps were first identified in *S. cerevisiae* where Arps were named on the decreasing order of their similarity with the yeast actin (Poch and Winsor, 1997). Arps 1-10 are found in yeast with Arps 7 & 9 specific to yeast. Arp11 is found in higher organisms, there is also a group of orphan Arps as classified by sequence analysis techniques by the Arpanno server (Muller *et al.*, 2005). Arp1 also known as cencentractin is involved in modulation of dynein motors. Arp2&3 form a seven member complex which is involved in filament formation like branching etc.. Arp 10 also has cytoskeletal function. The rest of the Arps 4-9 are probably involved in chromatin remodeling (Dayel and Mullins, 2004).

1.7 *D. discoideum* is an excellent model system to study the cytoskeleton

D. discoideum is a lower eukaryote whose cellular movement resembles the movement from higher organisms like leukocytes, satellite cells and others. It is a social amoeba, and, thus is a link between unicellular and multicellular organisms. It can live as a single cell or follow a developmental programme and form multicellular fruiting bodies. It harbors a barrage of actins and actin related proteins not found in any lower organism and hence is a prime attraction for cytoskeletal research. Developmental and molecular genetics tools have been well established and the genome is completely sequenced which favors molecular biology and bioinformatic research (Dictybase 2009) (Eichinger *et al.*, 2005) into the cytoskeleton. *D. discoideum* has an AT rich 34MB genome consisting of six chromosomes (Eichinger *et al.*, 2005). *D. discoideum*

has a short life cycle which makes it a good model to study signaling and development of the cytoskeletal system.

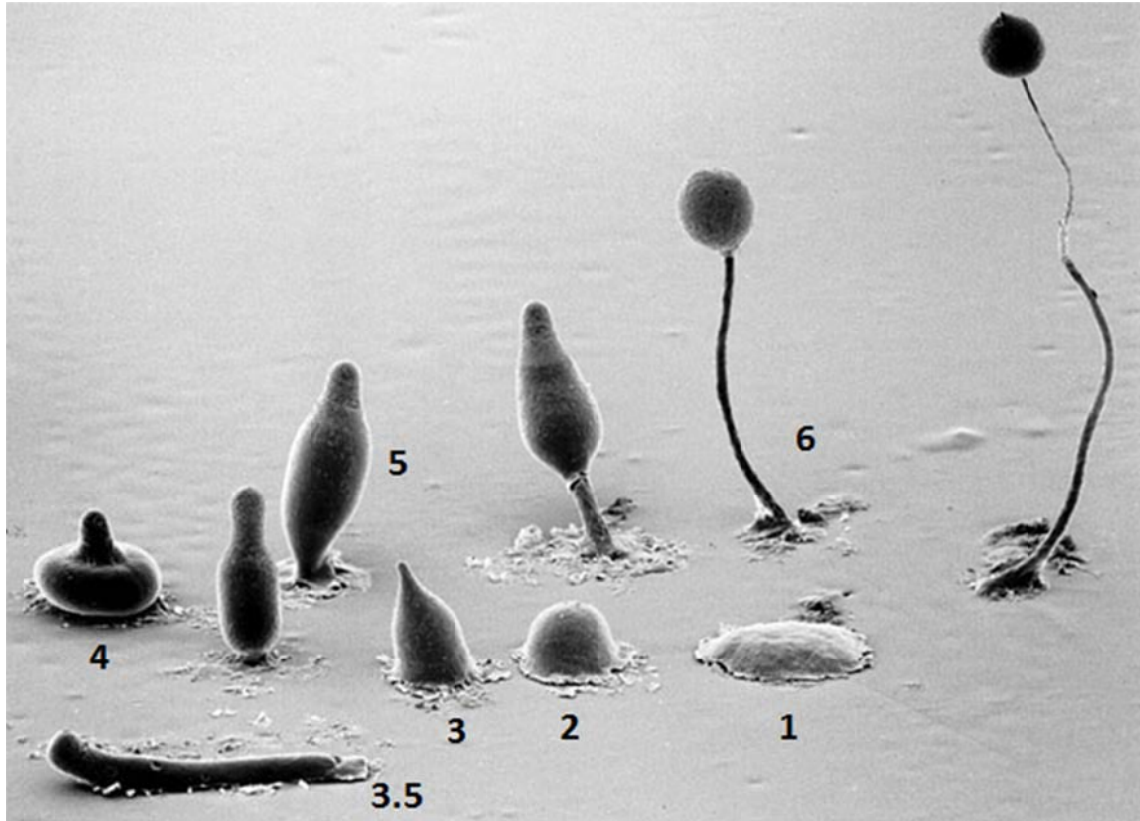


Figure 7: Scanning electron micrograph(Blanton) showing the different developmental stages of *D. discoideum*. Shown in the figure are; 1 loose aggregation, 2 tight aggregation, 3 finger, 3.5 slug, 4 mexican hat, 5 culmination, 6 fruiting body.

1.8 Aims of the project

D. discoideum has multitudes of actins isoforms. One of the aims of the project was therefore, to identify all the members of the *D. discoideum* actin superfamily (actinome) and to organize them using bioinformatics means.

In addition, the *D. discoideum* actinome should be compared with actinomes of other organisms. Selected Arps, actins and actin binding proteins were to be studied in detail by molecular, cell biological and biochemical approaches.

2. Materials and methods

2.1 Materials

Cell culture plates, 24 wells, flasks	Nunc
Cell culture dishes, 100 mm × 20 mm	Greiner bio-one
Cell culture dishes, 3.5 mm with glass bottom	MatTek Corporation
Dialysis membranes Type 8, 20, 27, 25A	Biomol
Gel drying membrane	Festata
Gel-blotting-paper	GB002
Nitrocellulose membrane Protran BA85	Whatman
PCR tubes 0.5 ml	Peqlab
Petri dishes 92 mm × 16 mm	Sarstedt
Pipettes, 10 und 25 ml	Nunc
15 ml and 50 ml tubes	Sarstedt
1.5 ml centrifuge tubes	Sarstedt
Pipette tips	Gilson
Plasmid DNA purification maxi kit	Qiagen
QIAprep spin miniprep and gel extraction kits	Qiagen

2.1.1 Software packages

Word processing:

- Word 2003: Microsoft inc,
- Acrobat reader 5.0: Adobe systems inc,
- Acrobat distiller 5.0 : Adobe systems inc.

Graphics and image processing:

- ImageJ 1.34n : Wayne rasband,
- Adobe photoshop 5.5: Adobe systems inc,
- Coreldraw 12 2003 : Corel corporation

Microscopy:

- AxioVS40 V4.3.101 : Carl Zeiss vision GmbH

Database:

- AskSam 4.0 : Seaside software

Sequence analysis:

- Bioedit: North Carolina state university

Sequence alignment:

- Clustalw: European bioinformatics institute

Bayesian tree construction:

- Mr.bayes: Florida state university

Tree graphics:

- Treeview: Taxonomy and systematics Glasgow

Molecular graphics:

- YASARA: YASARA biosciences GmbH,
- Chime: MDL information systems, inc

2.1.2 Servers

NCBI blast: (National Institute for Biotechnology Information) used for retrieving actinome members from the genomes of different organisms

<http://blast.ncbi.nlm.nih.gov/Blast.cgi>

pfam 22.0: (Wellcome Trust Sanger Institute) used for retrieving actinome members based on the presence of a specific actin profile

<http://www.sanger.ac.uk/Software/Pfam/>

Weblogo version 2.8.2: (University of California, Berkeley,) used for making sequence logos

<http://weblogo.berkeley.edu/>

Netnes 1.1: (Centre for biological sequence analysis, Denmark) mapping nuclear export signals

<http://www.cbs.dtu.dk/services/NetNES-1.1/>

PredictNLS (Jun 7, 2000) : (Rost lab, Munich) mapping nuclear localization signals

<http://www.predictprotein.org/> [PredictNLS]

<http://rostlab.org/cms/> [Rost lab, Munich]

Signalp 3.0: (Centre for Biological Sequence Analysis, Denmark) mapping the presence of export signals

<http://www.cbs.dtu.dk/services/SignalP/>

TMHMM 2.0: (Centre for Biological Sequence Analysis, Denmark) mapping the presence of transmembrane helices

<http://www.cbs.dtu.dk/services/TMHMM/>

PEST find: (European Bioinformatics Institute) identifying pest signals

<https://emb1.bcc.univie.ac.at/toolbox/pestfind/pestfind-analysis-webtool.htm>

Primer3: (Broad institute, White Head Institute) prediction of PCR primers

http://biotools.umassmed.edu/bioapps/primer3_www.cgi

Itol 1.3: (European Molecular Biology Laboratory) making circular phylogenetic trees

<http://itol.embl.de/>

2.1.3 Reagents

Unless otherwise mentioned all laboratory chemicals used were purchased from BioMol, Merck, Roth, Serva or Sigma and had the degree of purity 'p.a.'

Agar-Agar	Euler
Agarose	Biozym
Bacto-Peptone/-Tryptone	Oxoid
BCIP (5-Bromo-4-chloro-3-indolylphosphate-p-toluidinsalz)	Gerbu
Bovine serum albumin	Fraction
DTT (1,4-Dithio-D,L-threitol)	Gerbu
EGTA	Gerbu
IPTG (Isopropyl- β -D-thiogalactopyranoside)	Gerbu

Nucleotide	Peq labs
Proteose peptone	Oxoid

2.1.4 Centrifuges and rotors:

Centrifuges

GS-6KR Centrifuge, J2-21M/E Centrifuge, J6-HC Centrifuge:	Beckman
Table top centrifuge 5415:	Eppendorf

Rotors

JA-10, JA-14, JA-20, Ti 35, Ti 45, Ti 70, TLA 100.3:	Beckman
--	---------

2.1.5 Media

All media and buffers used were prepared with deionised water, which had been filtered over an ion exchanger (Millipore), and were sterilized either by autoclaving or by passing them through a micro filter.

AX-medium (pH 6.7)

14.3 g peptone
 7.15 g yeast extract
 50 mM glucose
 3.5 mM Na_2HPO_4
 3.5 mM KH_2PO_4
 made up to 1 l with H_2O

SM-agar plates (pH 6.5)

9 g agar
 10 g peptone
 50 mM glucose
 1 g yeast extract

4 mM MgSO_4

16 mM KH_2PO_4

5.7 mM K_2HPO_4

made up to 1 l with H_2O

Soerensen phosphate buffer (pH 6.0) (Malchow *et al.*, 1972)

14.6 mM KH_2PO_4 , pH adjusted with KOH

Phosphate agar plates (pH 6.0)

10 g Bacto-agar

2 mM Na_2HPO_4 buffer

HL5-medium (pH 7.5)

5 g yeast extract

10 g proteose peptone

50 mM glucose

8.5 mM KH_2PO_4

made up to 1 l with H_2O

PBS (10x, pH 6.5)

70 mM Na_2HPO_4

30 mM KH_2PO_4

150 mM NaCl

0.1% NaN_3

LB-medium for *E. coli* (pH 7.4)(Sambrook and Russell, 2001)

10 g bacto-tryptone

5 g yeast extract

86 mM NaCl

made up to 1 l with H₂O.

pH was adjusted with NaOH and for agar plates the medium was cooled to 40°C before the addition of antibiotics.

LB rich medium (pH 7.0)

20 g bacto-tryptone

10 g yeast extract

86 mM NaCl

made up to 1 l with H₂O.

5 ml of 0.5 M sodium phosphate was added prior to autoclaving

2.1.6 Buffers

10x Tris/Borate-buffer (TBE, pH 8.3)

890 mM Tris/HCl

890 mM boric acid

TE-buffer (pH 8.0)

10 mM Tris/HCl

2 mM EDTA

autoclaved before use

TEDABP-buffer (pH 8.0)

10 mM Tris/HCl

1 mM EGTA

1 mM DTT

0.02% NaN₃

prior to use the following ingredients were mixed with

1 mM benzamidine

0.5 mM PMSF

MEDABP-buffer (pH 6.5)

10 mM MES

1 mM EGTA

1 mM DTT

0.02% NaN₃

1 mM benzamidine

0.5 mM PMSF

IEDANBP-buffer (pH 7.6)

10 mM Imidazole

0.2 M NaCl

1 mM EGTA

1 mM DTT

0.02% NaN₃

prior to use the following ingredients were added:

1 mM benzamidine

0.5 mM PMSF

2.1.7 Instruments

Axiovert Microscopes 25, 35, M200

Zeiss

Dounce homogenisor

Kika

Eagle Eye II

Stratagene

Electroporator

BioRad

Heating block/Shaking Thermomixer 5436

Eppendorf

PCR-Thermocycler Uno

Biometra

pH-Meter

pH526 WTW

Photometer Ultrospec 2100 pro

Amersham

Protein transfer Trans-Blot SD

BioRad

Quartz cuvettes

Starna

Superdex 200, Superose 6 / 12 10/300 GL

GE

Ultrasonicator 820/H :

Elma

UV-transilluminator :	Stratagene
Vortex :	Bender & Hobein
Waterbath :	GFL.

2.1.8 Microbial strains

***E.coli* strains:**

DH5 α

JM 105

BL21 RIL Stratagene

Klebsiella aerogenes(Williams and Newell, 1976)

***D. discoideum* strains:**

AX2-214 (Laboratory wild type)

2.1.9 Vectors

pGEX 6P-1 GE. It has a tac promoter for chemical inducibility and high-level expression of GST-tagged recombinant proteins. An internal lacIq gene for use in any *E. coli* host.

pLPBLP(Faix *et al.*, 2004). The gene targeting vector pLPBLP has a *loxP* sites in the same orientation flanking both sides of the Bsr expression cassette *act15/Bsr*. An oligonucleotide cassette has translational stop codons in all six reading frames. The restriction enzyme sites outside of the floxed-*Bsr* cassette, such as SmaI, permits the cloning of 5' and 3' gene sequences for targeted disruption.

2.2 Molecular Biological Methods

2.2.1 Isolation of *D. discoideum* genomic DNA using a high pure PCR template preparation kit (Roche)

D. discoideum cells from Petri dishes or shaking cultures were harvested and washed with Soerensen buffer. The pellet was resuspended in 200 µl of PBS; 200 µl of binding buffer; 40 µl of proteinase K (20 mg/ml) were added immediately to the cell suspension and the sample incubated at 72°C for 10 min. 100 µl of isopropanol were added to the suspension which was then transferred into the upper reservoir of the filter set and centrifuged for 1 min at 8,000 g. The flowthrough was discarded. The filter tube was washed once with inhibitor removal buffer and twice with wash buffer (1 min / 8,000 g). Residual wash buffer was removed from the filter by centrifuging the empty filter tube for an additional minute at 8,000 g. The filter tube was transferred into a clean 1.5 ml Eppendorf tube and the DNA was eluted with 200 µl prewarmed (70°C) elution buffer by spinning the sample for 1 min at 8,000 g. The genomic DNA was stored at -20°C.

2.2.2 Polymerase chain reaction (PCR)

The amplification of DNA fragments and generation of site-directed mutagenesis was carried out by polymerase chain reactions. Taq polymerase (Roche), Taq buffer, dNTPs were used as usual. The reaction was performed on a thermocycler (Biometra UNO) in either a 50 µl or a 100 µl reaction. cDNA or genomic DNA were used as a template. A brief denaturation step was always done before the start of the actual 25-35 cycles for amplification. Annealing temperatures were calculated according to the formula of Suggs *et al* (Suggs *et al.*, 1981)

$$4(N_G + N_C) + 2(N_A + N_T) - 10 = T_a (^{\circ} C)$$

N = the number of Adenine (N_A), Thymidine (N_T), Guanidine (N_G) or Cytosine (N_C) in oligonucleotides.

Reaction conditions in most assays:

Denaturation: 95°C for 120s

Denaturation: 95°C for 30s

Annealing: variable temperature for 60s

Elongation: 69°C for 1 min for every 1000 bp

Final extension 70°C for 10 min

No. of cycles: 30

The annealing temperature is dependent of the primer sequences.

10x PCR buffer

200 mM Tris/HCl (pH 8.8)

100 mM KCl

100 mM $(\text{NH}_4)_2\text{SO}_4$

20 mM MgSO_4

1% Triton X-100

1 mg/ml BSA (Nuclease-free)

2.2.3 Agarose gel electrophoresis(Sambrook and Russell, 2001)

Electrophoretic separation of DNA was carried out using 0.7% (w/v) agarose gels prepared in 1 x Tris-borate EDTA buffer at 1-5 V/cm. Running and gel buffer were identical. The gels were cast in chambers of various sizes (4 x 7 to 20 x 20 cm), the DNA to be separated was mixed with 0.2 vol of DNA sample buffer and loaded onto the gels. At the end of the run, the gel was stained in a solution of ethidium bromide (5 µg/ml) for 10-30 min, followed by destaining in water for about 30 min. Subsequently the gels could be observed under UV light at 312 nm and photos taken with the Eagle Eye CCD camera system (Stratagene, Heidelberg).

DNA sample buffer

40% sucrose

0.5% SDS

0.25% bromophenol blue

-Taken up in TE buffer

2.2.4 Determination of DNA concentration

DNA was diluted 100 times in sterile H₂O and the absorbance was measured at 260 nm and 280 nm. An O.D₂₆₀ of 1.0 corresponds to 50 µg/ml of double stranded DNA. An A_{260:280} ratio of 1.80 indicated highly pure DNA free from contaminating protein. Alternatively, the DNA concentration could also be estimated by comparing the intensity of bands after agarose gel electrophoresis with the bands in the Mass Ruler lane.

2.2.5 Isolation of DNA fragments from agarose gels

Extraction and purification of DNA from agarose gels in Tris-borate buffer were performed using the QIAquick gel extraction kit (Qiagen) according to the manufacturer's instructions. The kit provides buffers, and all centrifugation steps were carried out at 14,000 rpm at room temperature using a tabletop microcentrifuge. The DNA was excised from the agarose gel with a clean scalpel, the gel slice weighed, and 3 vol buffer QG were added to 1 vol gel (100 mg~100 µl). The tube was then incubated in a heating block at 50°C for 10 min or until the gel slice was entirely dissolved. To increase the yield of DNA fragments (< 0.5kb and > 4kb), 1 gel vol isopropanol was added to the sample and mixed. The sample was then loaded onto a QIAquick spin column (Qiagen) and centrifuged for 1 min in order to bind DNA to the resin. The flowthrough was collected in a 2 ml collection tube and discarded. For washing, 750 µl of buffer PE was added to the column and spun for 1 min. The flowthrough was discarded and the column was centrifuged for an additional 1 min to remove residual ethanol completely. The column was then placed into a clean 1.5 ml tube and the DNA was eluted by the addition of 30-50 µl of elution buffer (10 mM Tris/HCl, pH 8.5) or dH₂O to the column followed by a centrifugation for 1 min.

2.2.6 Purification of PCR products

For cloning purposes, the PCR products were purified using the QIAquick PCR Purification Kit from Qiagen following the manufacturer's protocol. The kit-provided buffers were used and all centrifugation steps were done at 14,000 rpm at room temperature using a table top microcentrifuge. 5 vol of PB buffer were added to 1 vol

of the PCR product and mixed. The mixture was applied to a QIAquick spin column and centrifuged for 1 min to bind DNA to the column while the flowthrough collected in a 2 ml collection tube was discarded. DNA was washed with 750 µl of PE buffer by a centrifugation for 1 min. Residual ethanol was removed by an additional centrifugation for 1 min. The spin column was then placed into a clean 1.5 ml tube and the DNA was eluted with 50 µl of EB buffer. The purified DNA was subsequently used for restriction digestion

2.2.7 DNA cleavage with restriction enzymes

For cloning purposes usually 1 µg of DNA (plasmid or PCR product) was digested in a 50 µl reaction with 10 U of each enzyme. The buffer suitable for the restriction enzyme(s) was added and the volume was adjusted with dH₂O. The tube was incubated at the 37°C overnight. The digestion of the vector was monitored on an agarose gel. For screening of putative clones, the reaction tubes were incubated at 37°C for 45-60 min with 0.1-0.25 U of enzyme / 2 µg of DNA.

2.2.8 Mini preparation of plasmid DNA using Qiagen miniprep kit

The plasmid DNA from *E. coli* was isolated with the QIAprep Spin Miniprep Kit from Qiagen. All buffers were supplied with the kit. Single colonies from LB plates were inoculated and grown in LB medium with appropriate selection pressure. Cells were pelleted and resuspended in 250 µl buffer P1 containing RNase. The cells were lysed by adding 250 µl alkaline lysis buffer P2 and the suspension was mixed gently by inverting the tube 4-6 times. The suspension was then neutralized by mixing the suspension with the addition of 350 µl buffer N3. The mixture was centrifuged (10 min / 10,000 g). The plasmid containing supernatant was passed through the QIAprep spin column to allow binding to the resin. The column was washed with 750 µl buffer PE, and centrifuged for an additional minute to remove residual wash buffer. The QIAprep column was then placed into a clean 1.5 ml tube and DNA was eluted with 50 µl of EB buffer.

2.2.9 Mini preparation of plasmid DNA by the method of Holmes and Quigley

Overnight cultures of *E. coli* in LB medium (with ampicillin; 10 µg/ml) were spun and the pellet was resuspended in 200 µl of STET/lysozyme buffer. The suspension was boiled in a water bath for 1 min, centrifuged (13,000 rpm, 10 min, room temperature) and the insoluble cell debris was removed using a sterile tooth pick. The nucleic acids in the supernatant were precipitated with 200 µl of isopropanol for 5 min at room temperature and sedimented using a table top centrifuge (13,000 rpm, 30 min, room temperature). The DNA pellet was washed with 70% ethanol, air dried and the pellet was resuspended in 50-100 µl sterile H₂O or in TE (Holmes and Quigley, 1981).

2.2.10 Phosphatase treatment(Sambrook and Russell, 2001)

In order to prevent self-ligation of a linearised vector, the phosphate group at the 5' end was removed using alkaline phosphatase from calf intestine. Dephosphorylation was performed in a 50 µl reaction volume, whereby 1-5 µg of linearised DNA vector was incubated with 1-2 U of alkaline phosphatase in 1 x phosphatase buffer at 37°C for 1 h. To inactivate the enzyme, the DNA was purified over a Qiagen column as described above.

2.2.11 Ligation of DNA into a plasmid vector

Ligation of CIP treated vector and insert was carried out overnight at 16°C using 1.5 U of enzyme, in a minimum volume (usually not more than 30 µl). The vector and the insert were mixed at a molar ratio of 1:3.

2.2.12 Preparation of chemically competent cells

250 ml LB medium were inoculated with an over night preculture, and grown in a 37°C shaker at 220 rpm until an OD₆₀₀ of 0.4-0.6 was reached. Cells were put on ice for 30 min and then centrifuged for 20 min (4°C, 4,000 g). Cells were washed twice with ice-cold CaCl₂/glycerol. Finally the pellet was resuspended in 2 ml of CaCl₂ / glycerol buffer and kept on ice in 200 µl aliquots for 30 minutes, cold shock-frozen in liquid nitrogen and stored at -70°C.

Transformation buffer (CaCl_2 /glycerol)

60 mM CaCl_2

15% glycerol

10 mM PIPES

pH was adjusted to 7.0

2.2.13 Transformation of *E. coli*

For transformation, competent *E. coli* cells were thawed on ice and incubated with the ligation mix on ice for 20 min followed by a heat shock at 42°C for 90 s. The cells were immediately placed on ice for 1 min and revived with LB medium. Cells were kept at 37°C with agitation for 45-60 min and plated on LB plates containing the appropriate antibiotics. Plates were incubated at 37°C overnight to obtain *E. coli* colonies harbouring the transformed plasmid. Single clones were inoculated into tubes containing 2-5 ml LB with antibiotics. Cells were grown on a 37°C shaker for 7 h or overnight. The positive transformants were detected by mini preparation of DNA plasmid.

2.2.14 *E. coli* permanent cultures

Overnight grown cultures of bacteria were centrifuged and the pellets were resuspended in fresh LB medium containing 7% DMSO. The cells were cold shock frozen in liquid nitrogen and stored at -70°C.

2.2.15 Maxi preparation of plasmid DNA

To isolate large quantities of plasmid DNA, DNA was isolated from 250-400 ml cultures using a DNA Maxi-kit. The cells were harvested by centrifugation (4,000 g; 5 min; 4°C). The cell pellet was resuspended in 12 ml of buffer S1+RNase A, lysed by adding 12 ml of buffer S2 and mixed gently by inverting the vial 6-8 times. The mixture was incubated at room-temperature for not more than 5 min. The lysate was neutralized by carefully adding 12 ml of pre-cooled buffer S3, incubated on ice for 15 min, transferred to the centrifugation tubes and spun (12,000 g, 40 min, and 4°C). The supernatant was passed through the NucleoBond AX 500 column that was pre-

equilibrated with 6 ml of buffer N2. The column was washed with 35 ml of wash buffer. The plasmid DNA was eluted with 15 ml of buffer N5 and precipitated with 11 ml of isopropanol at room temperature. DNA was pelleted by centrifugation at 15,000 g, 4°C, for 30 min. The DNA pellet was washed with 5 ml of 70% ethanol (centrifuged at 15,000g; room temperature; 10 min), air-dried and dissolved in 500 µl 10 mM Tris. pH 8.0

Buffer S1

50 mM Tris-HCl

Buffer S2

200 mM NaOH

10 mM EDTA 1% SDS

100 ug/ml RNase A

Buffer S3

2.8 M potassium acetate

Buffer N2

100 mM Tris pH 5.1

15% ethanol

900 mM KCl

0.15% Triton X-100

pH adjusted to 6.3 with H_3PO_4

Buffer N3:

100 mM Tris

15% ethanol

1.15 M KCl

pH adjusted to 8.5 with H_3PO_4

2.3 Biochemical methods**2.3.1 SDS-polyacrylamide gel electrophoresis (SDS-PAGE)**

Protein mixtures were separated by discontinuous SDS-PAGE(Laemmli, 1970). For this purpose, 10, 12 and 15% acrylamide resolving gels with 3% stacking gels were used (7.5 x 10 x 0.05 cm). The stacking gel deposits the polypeptides to the same

starting level at the surface of the resolving gel, and subsequently the SDS-polypeptide complexes are separated in the resolving gel according to size under uniform voltage and pH. Prior to SDS-PAGE, 3 x SDS gel loading buffer was added to the protein samples to be separated. The mixture was boiled for 3-5 min. Electrophoresis was carried out at a constant voltage of 150 V, after which the gel could be stained with Coomassie blue dye and destained for direct observation of the protein bands. Proteins from the gel could also be blotted onto a nitrocellulose membrane and detected indirectly via antibodies. As standard, a mixture of proteins of defined molecular masses was electrophoresed.

3x SDS-PAGE sample buffer

150 mM Tris/HCl / 1.9 M glycine (pH 6.8)

30% Glycerol

6% ($\frac{w}{v}$) SDS

15% β -mercaptoethanol

0.3% bromophenol blue

2.3.2 Coomassie blue staining of proteins

Following SDS-PAGE, gels were stained in Coomassie blue solution by shaking for 15-30 min, after which the unbound dye was removed by shaking in a destaining in 10% ethanol / 7% acetic acid.

Coomassie blue solution

0.1% Coomassie brilliant blue R 250

50% methanol

12% acetic acid

Solution filtered through a Whatman filter

2.3.3 Western blotting:

Following separation of protein by SDS-PAGE, the proteins were transferred from gels onto nitrocellulose membranes (Whatman BA85) according to a modified protocol of Towbin *et al* (Towbin *et al.*, 1979) by the means of a protein transfer

apparatus (Trans-Blot SD, BioRad). In this “semi-dry” method, the gel and its attached nitrocellulose membrane were sandwiched between several pieces of Whatman 3MM filter paper which had been soaked in transfer buffer, and protein transfer was carried out at room temperature at 15 V for 45 min, after which the nitrocellulose filter was blocked for at least 1 h or overnight at room temperature in 5% (w/v) milk powder in NCP buffer. The membrane was then washed five times with NCP. The membrane was incubated with primary antibody diluted in 0.1% BSA solution in NCP overnight or at least for 3 h. After washing 3 times with NCP buffer, the membrane was incubated with a secondary antibody conjugated with alkaline phosphatase (diluted 1:2,000 in NCP buffer + 1% BSA) for at least 3 h. The membrane was washed 5 times (15 min each) with NCP and once with 0.1 M Na_2CO_3 -buffer, pH 10.2. Finally, the protein bands were detected by using 25 ml of 0.2 mg/ml BCIP diluted in 0.1 M Na_2CO_3 .

Transfer buffer (pH 8.5):

25 mM Tris/HCl

190 mM glycine

20% methanol

20x NCP buffer (pH 7.2)

200 mM Tris/HCl

3 M NaCl

20 ml Tween 20

0.02% SDS

2% NaN₃

2.3.4 Expression of GST-tagged protein

Constructs in pGEX vectors were transformed into DH5 α cells (Novagen). Cultures were inoculated and grown overnight at 37°C, diluted tenfold and grown at 21°C to an OD₆₀₀ of 0.4 - 0.8. Expression was induced with usually 0.5 mM IPTG and cells were grown overnight at 21°C. After harvesting and washing following routine procedures, cells were resuspended in lysis buffer, opened by sonication in the presence of 0.5

mg/ml lysozyme. Lysates were centrifuged at 10,000 g for 45-60 min at 4°C and the supernatants were coupled to Glutathione-Sepharose 4B (Sigma) by recycling for 3-4 h. The matrix was washed with 10-20 column volumes of lysis buffer and the bound proteins were eluted with lysis buffer containing 20 mM reduced glutathione. The presence of protein in different fractions was tested by Bradford's method (Bradford, 1976). Protein containing fractions were analyzed on SDS-PAGE. The appropriate fractions were pooled and dialysed against PBS.

Lysis Buffer

1 x PBS

1 mM DTT

5 mM benzamidine

1 mM PMSF

Protease inhibitor cocktail (Sigma)

2.4 Cell Biological Methods

2.4.1 Growth in liquid medium and on agar plates

Wild type AX2 strain was cultured axenically in either AX medium or in HL5 medium in Erlenmayer flasks. The media were supplemented with the appropriate antibiotics when cultivating the mutants. The generation time for the wild type at 21°C and 150 rpm is about 8-10 h. Cells were also cultured on Petri dishes. For this purpose AX2 wild type cells were suspended in 12 ml HL5 medium. For all cell biological studies, cells were allowed to grow maximally up to a cell density of 5×10^6 cells/ml to avoid the stationary phase. For large scale preparations cells were cultivated in 4 x 2.5 l cultures up to a cell density between 5×10^6 to 1.2×10^7 cells/ml.

2.4.2 Preservation of spores

Cells from liquid culture were harvested by centrifugation at 1,000 g for 10 min, washed once with cold Soerensen buffer, and resuspended at a cell density of about 2×10^8 cells/ml. 500 µl of the cell suspension were spread out per phosphate agar plate. The cells were able to develop into fruiting bodies within 2-3 days. The spores were then harvested by knocking onto the lid of the Petri dish and taken up in 10 ml of cold

Soerensen buffer. 1 ml aliquots were dispensed into Nunc tubes (2.2 ml), shock-frozen in liquid nitrogen and stored at -70°C . For inoculation with spores, an aliquot was thawed at room temperature and cultured either in AX or in HL5 medium, whereby after 3 days at 21°C and 150 rpm, a cell density of about 5×10^6 cells/ml was reached.

2.4.3 Storage of *D. discoideum* cells at -70°C

For preservation of *D. discoideum* cells, axenic cultures were harvested and resuspended at a cell density of about 5×10^7 cells/ml in ice-cold freezing medium, (1.1 ml containing 900 μl of HL5 or AX medium, 100 μl of horse serum and 100 μl of DMSO). Cells were distributed as 1.1 ml aliquots into sterile tubes, pre-cooled on ice. The tubes were incubated at 4°C for one hour, -20°C for 2-4 h and then stored at -70°C .

2.4.4 Transformation of *D. discoideum* cells

D. discoideum cells were transformed with the appropriate plasmids by electroporation. Approximately 10^7 growth-phase cells were washed at 0°C - 4°C , twice in 17 mM K-phosphate buffer, pH 6.0 and once in electroporation buffer (EP; 50 mM sucrose, 10 mM Na-K-phosphate buffer, pH 6.1). The cell pellet was then resuspended in 700 μl EP and the suspension gently mixed with ~ 35 μg of DNA at 0°C in a 4 mm electroporation cuvette. Electroporation was performed for 1 ms, at 1 kV and 10 μF with the Gene Pulser Xcell (Biorad); two pulses were applied at 5 s intervals. After electroporation, the cells were transferred to a 9 cm plastic Petri dish and shaken at ~ 40 rpm for 15 min at 20°C . The suspension was adjusted to 2 mM CaCl_2 and 2 mM MgCl_2 and shaking was continued for another 15 min. Finally, 12 ml axenic growth medium was added, and the cells were allowed to recover for 24 h. After the 24 h recovery period, the cells were placed in axenic medium containing the appropriate selection pressure to select for transformants.

Electroporation buffer (pH 6.1)

50 mM sucrose

pH was adjusted to 6.1 with KOH solution, and the buffer was sterilized by filtration

2.4.5 Cloning of transformants

Following 10-15 days of selection, the transformants were washed from the Petri dishes, diluted and plated together with an aliquot of *K. aerogenes* suspension onto SM agar plates. The transformants were diluted to about 50-100 cells per plate for better isolation of single colonies. Within 2-3 days at 21°C, round clearing plaques were visible. These were picked with sterile toothpicks and cultured in 24-well plates in HL5 medium containing selection pressure for transformants (G418; with the dilution of 1:1,000) and antibiotics mixture (ampicillin/streptomycin) to remove the contaminating bacteria. The axenic cultures were then grown in larger quantities, harvested and further analyzed.

2.4.6 Affinity purification

The antibody was subjected to affinity purification prior to immunofluorescence. Nitrocellulose membrane was cut in to small squares and soaked in NC-buffer. Up to 1 mg of pure antigen was added and the membrane incubated at 4°C for 1-2 hrs. The membrane was washed 5 times with 1x PBS, followed by 1 hr incubation with 1x PBS and 1% BSA. It was washed again 5 times with NCB. 2.5 ml of antiserum were added and incubated at 4°C over night. Followed by 5 washing steps with NCB, and then one wash with water. The antibody was eluted with 0.1 M acetic acid in 150 mM NaCl. The eluent was neutralized in a new Eppendorf tube with 40 µl of 1M Tris, and finally dialyzed in 1x PBS overnight at 4° C.

2.4.7 Indirect immunofluorescence

Studies of subcellular localisation were performed via indirect immunofluorescence. For this assay, coverslips to be used were washed with 3.6% HCl followed by dH₂O. Exponentially growing *D. discoideum* cells were harvested, washed twice with Soerensen buffer, and 1×10^6 cells were allowed to attach to the coverslips for 15 min, after which excess fluid was removed and the cells were fixed in cold methanol for 10 min at -20°C followed by 30 min air drying. Alternatively, cells could be fixed with paraformaldehyde/picric acid solution (2% paraformaldehyde, 10 mM PIPES, 15% saturated picric acid, pH 6.0) for 15 min and then washed several times with

PBS/glycine and PBG. The cells were then permeabilised with 70% ethanol for 5 min at room temperature and washed with PBS/glycine followed by PBG.

After fixation, cells were incubated with undiluted hybridoma supernatants for at least 2 h before being washed with PBG and then subjected to 1 h incubation with fluorescently labelled goat anti-mouse IgG (diluted in PBG). F-actin was labelled either with TRITC-labelled phalloidin or a monoclonal antibody against *D. discoideum* actin (Simpson *et al.*, 1984). Nuclei were stained either with DAPI (4, 6-diamidino-2-phenylindole, Sigma) or TO-PRO diluted in PBG. After incubation cells were washed several times with PBG, PBS and briefly with dH₂O before being embedded in gelvatol and kept at 4°C overnight.

10x PBS (for immunofluorescence, pH 7.4)

1.37 M NaCl

0.027 M KCl

0.081 M Na₂HPO₄ in 1x PBS, sterilized by passing through a 0.5 µm filter

PBG

0.5% BSA

0.05% fish gelatine

0.015 M KHPO₄

3. Results

3.1 The *D. discoideum* actinome

A multitude of actin hits were observed when the *D. discoideum* genome was blast searched for actins. To get a clear picture a robust technique was needed for the identification of the members of the actinome. We adopted the profile based 'Hidden Markov models' used in the Pfam (protein family database)(Finn *et al.*, 2006) to identify the members of the *D. discoideum* actinome(Joseph *et al.*, 2008). 33 genes which coded for the corresponding proteins with an actin domain profile, were included into the actinome (Table 1). Seven genes failed to meet the criteria by coding for just partial actin domains (Table 2). The derived amino acid sequences were also included in the actinome but the corresponding DNA regions were classified as pseudogenes. These genes would give rise to partial domains that range from 41 to 122 amino acids. In one case there is a complete gene but the coding sequence contains a frame-shift which would result in a miscoded protein.

Among the 33 derived proteins found to have the actin domain profile, 17 are identical to each other (Act8 group) and 17 of them are non-identical. The differences between the non-identical actins are small, e.g. only one D2E substitution in Act10, rather drastic, e.g. 295 non-identical amino acids in Act33. We identified three new actins in the *D. discoideum* actinome (DDB0234013, DDB0234014, DDB0234012) that have not been reported before and named the genes *act31*, *act32*, *act33*, respectively.

Table 1: The major members of the actinome (Joseph *et al.*, 2008). It includes the identical actins, non-identical actins and Arps. Act8 is the representative sequence for all of the identical actins.

Proteins with complete actin domain profile				
#	Protein	Gene	dictyBase ID	Swiss-Prot ID
1	Act3	<i>act3</i>	DDB0220458	P07829
2	Act8*	<i>act8</i>	DDB0216213	P07830
3	Act10	<i>act10</i>	DDB0220457	Q54GX7
4	Act17	<i>act17</i>	DDB0185125	Q554S6
5	Act18	<i>act18</i>	DDB0220459	P07828
6	Act22	<i>act22</i>	DDB0220460	Q553U6
7	Act23	<i>act23</i>	DDB0220461	Q55EU6
8	Act24	<i>act24</i>	DDB0220462	Q54HF1
9	Act25	<i>act25</i>	DDB0220463	Q54HF0
10	Act26	<i>act26</i>	DDB0220464	Q55CU2
11	Act27	<i>act27</i>	DDB0229353	Q54HE9
12	Act28	<i>act28</i>	DDB0229354	Q54HE7
13	Act29	<i>act29</i>	DDB0229355	Q54L54
14	Act31	<i>act31</i>	DDB0234013	Q55DY5
15	Act32	<i>act32</i>	DDB0234014	Q55DS6
16	Act33	<i>act33</i>	DDB0234012	Q54JL1
17	Filactin	<i>fia</i>	DDB0220465	Q54PQ2
18	Arp1	<i>arpA</i>	DDB0220489	Q54I79
19	Arp2	<i>arpB</i>	DDB0185179	Q96621
20	Arp3	<i>arpC</i>	DDB0219936	P42528
21	Arp4	<i>arpD</i>	DDB0233063	Q54UQ7
22	Arp5	<i>arpE</i>	DDB0234009	Q54E71
23	Arp6	<i>arpF</i>	DDB0234010	Q54KZ7
24	Arp8	<i>arpG</i>	DDB0234011	Q54JV5
25	Arp11	<i>arpH</i>	DDB0233828	Q54JY2

Identical actins of the Act8 group		
1	Act1	DDB0220444
2	Act2	DDB0185124
3	Act4	DDB0220448
4	Act5	DDB0220447
5	Act6	DDB0185126
6	Act7	DDB0220445
7	Act8	DDB0216213
8	Act9	DDB0220456
9	Act11	DDB0220449
10	Act12	DDB0216214
11	Act13	DDB0220454
12	Act14	DDB0220455
13	Act15	DDB0185015
14	Act16	DDB0185127
15	Act19	DDB0220446
16	Act20	DDB0220450
17	Act21	DDB0220451

Table 2: A list of actin pseudogenes (Joseph *et al.*, 2008). The Dictybase identification numbers were supplemented with names for these genes that we suggested to dictybase. The number of amino acids for each pseudogene is also tabulated.

Sl.No	Dictybase identification number	Gene name	Number of amino acids
1	DDB0237409	(<i>act30_ps</i>)	106
2	DDB0237453	(<i>act31_ps</i>)	41

3	DDB0237452	(<i>act32_ps</i>)	111
4	DDB0237450	(<i>act33_ps</i>)	135
5	DDB0237454	(<i>act34_ps</i>)	67
6	DDB0237455	(<i>act35_ps</i>)	88

3.2 Genomic locations of selected actin genes

As mentioned earlier there are 17 actin genes that are identical at the protein sequence level, but they differ at the nucleotide sequence level. Since these identical actins arise as paralogues by multiple duplications that occur at the gene level and not at the level of the whole genome(Hennig, 2004), it is significant to map the locations of these genes across the *D. discoideum* chromosomes. They are mapped and tabulated as shown in Table 3.

Table 3: The chromosomal locations of the identical actins(Joseph *et al.*, 2008). The corresponding strand is also mentioned.

Sl.No	Proteins identical to Act15	Chromosomal location
1	Act1	Chromosome 5 coordinates 2938326 to 2939456, Crick strand
2	Act2	Chromosome 2 coordinates 4708449 to 4709579, Watson strand
3	Act4	Chromosome 5 coordinates 2234318 to 2235448, Crick strand
4	Act5	Chromosome 5 coordinates 3186992 to 3188122, Crick strand

5	Act6	Chromosome 2 coordinates 4042347 to 4043477, Watson strand
6	Act7	Chromosome 3 coordinates 3684122 to 3685252, Watson strand
7	Act8*	Chromosome 1 coordinates 4602557 to 4603687, Watson strand
8	Act9	Chromosome 2 coordinates 4441895 to 4443025, Watson strand
9	Act11	Chromosome 5 coordinates 2167534 to 2168664, Watson strand
10	Act12	Chromosome 2 coordinates 4705207 to 4706337, Crick strand
11	Act13	Chromosome 2 coordinates 4481524 to 4482654, Crick strand
12	Act14	Chromosome 2 coordinates 4404974 to 4406104, Watson strand
13	Act15	Chromosome 2 coordinates 1767622 to 1768752, Watson strand
14	Act16	Chromosome 2 coordinates 1594273 to 1595403, Crick strand
15	Act19	Chromosome 2 coordinates 4040120 to 4041250, Watson strand
16	Act20	Chromosome 2 coordinates 4034200 to 4035330, Crick strand
17	Act21	Chromosome 2 coordinates 4038493 to 4039623, Crick strand

Chromosomes 2 and 5 were studied in detail with respect to the localization of the members of the actinome in these two chromosomes. Figure 8 shows the clustering of the members of the actinome in these two chromosomes, which suggests a wave of duplication events, confirmed by a phylogenetic assay shown later.

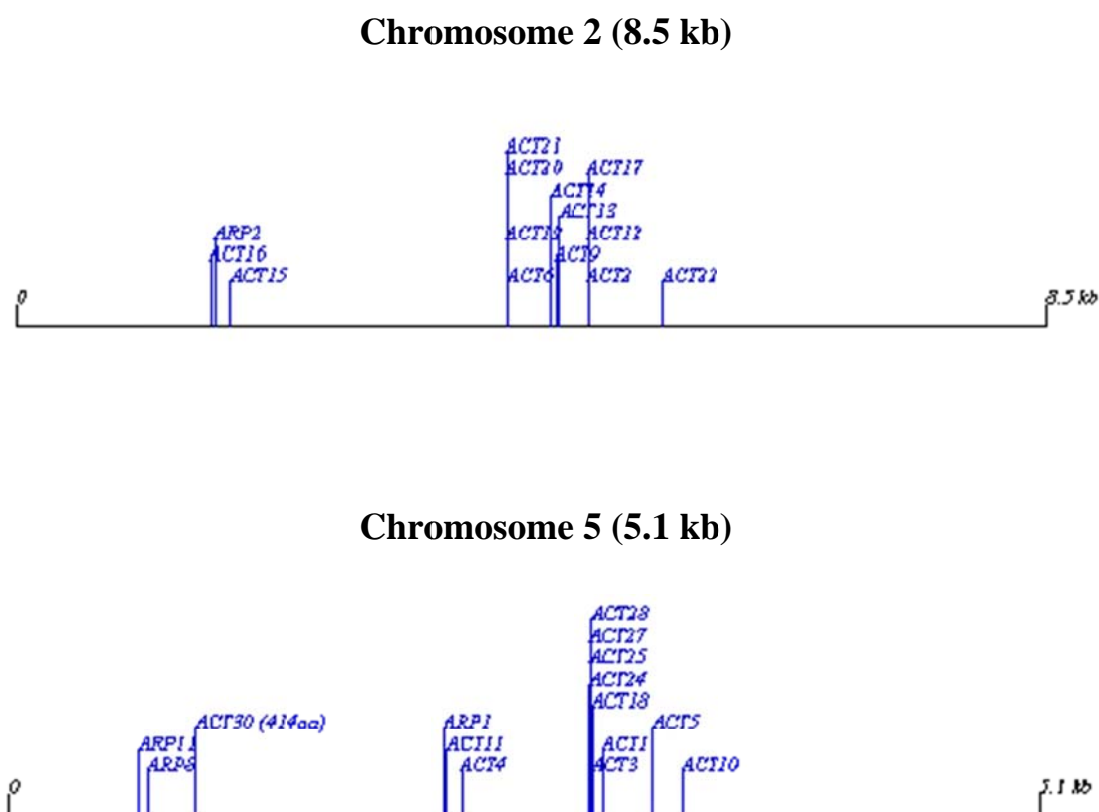


Figure 8: The corresponding locations of the actin genes on chromosomes 2 and 5. The size of each chromosome is also shown. Clustering of genes suggests repeated gene duplications.

3.3 Phylogeny of the *D. discoideum* actinome

To study the phylogeny of the *D. discoideum* actinome, a phylogenetic assay was performed which involves all the members of the actinome with the identical actins represented only once by Act8. The phylogenetic tree of the *D. discoideum* actinome shows Arp4 as closest to the bacterial actin-like protein MreB, which was used as an outgroup and is thought to be a putative ancestor of all actins. The most closely related actin is Act22 that differs from this group by three amino acid exchanges (A236S,

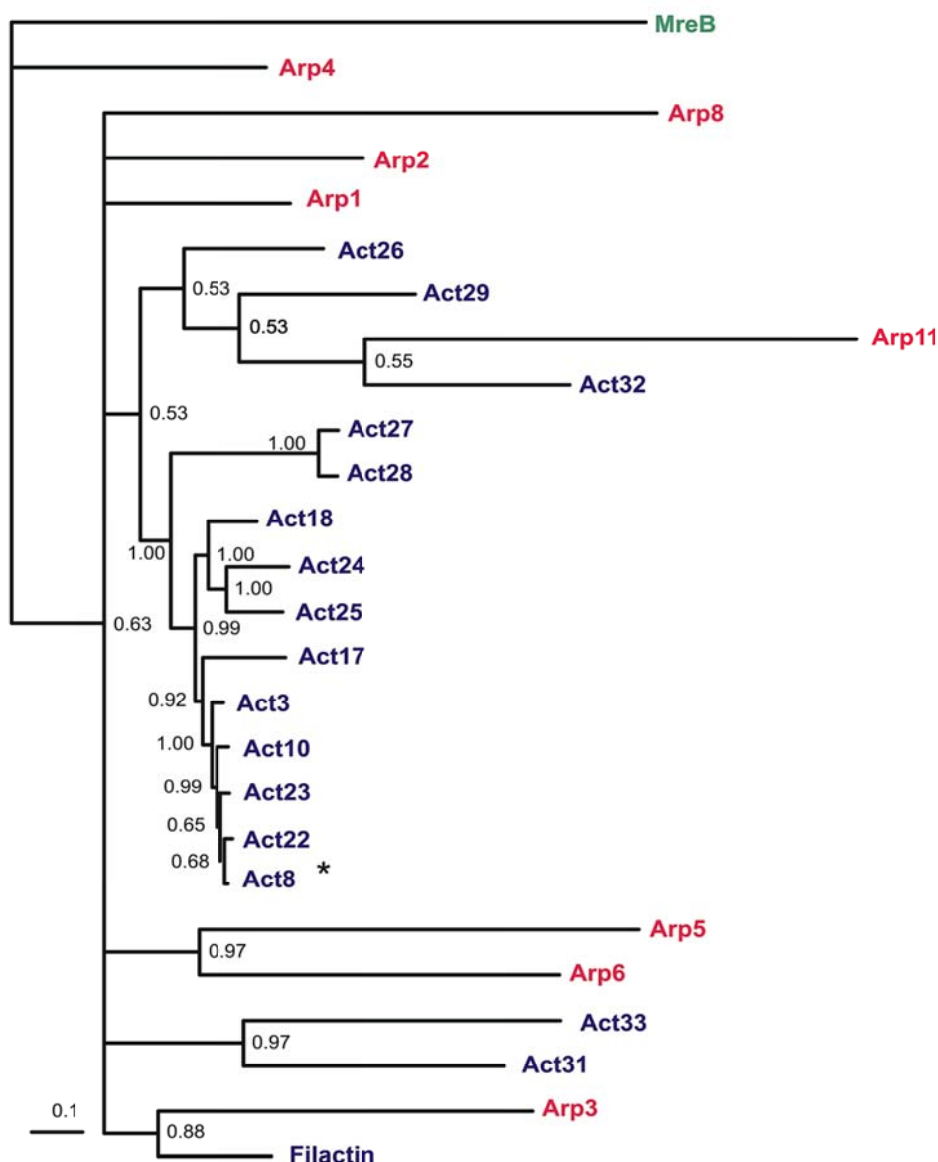


Figure 9: The evolutionary relatedness of the members of the actinome can be inferred from this tree (Joseph *et al.*, 2008). MreB (Figge *et al.*, 2004; Kruse and Gerdes, 2005), a prokaryotic actin homolog, forms the outgroup. Especially interesting are the locations of Arp4 and Act8. Arp4 is supposedly the oldest of the actin related proteins and is found closer to the outgroup. The pattern of duplication events that eventually lead to Act8 is interesting because of the paralogous duplication events that led to the Act17 followed by Act3, Act10, Act23, and the final gene level duplication that led to the formation of the two genes Act22 and Act8.

Y280F, A320S). Act10 with one single residue exchange only (D2E) is more distant, which reflects the scores in the permutation matrix used by the alignment program.

Exchanges from A/S, Y/F, and D/E score 1, 0, and 3 points, respectively, thus listing Act22 as more closely related to the 17 identical actins than Act10. An alignment of all Arps with the Act8 protein sequence is shown in Fig. 10. Filactin has not been included because it belongs to the bona fide actins and contains a compact actin domain that is highly homologous to conventional actin.

Eleven of the actin genes that code for the Act8 group are located on chromosome 2, four on chromosome 5, and one each on chromosomes 1 and 3. A phylogenetic tree of the DNA sequences from the identical actin proteins clearly shows the series of multiplication events (Fig.11, below). Most of the genes are clustered. This suggests a wave of gene duplications especially on chromosome 2. Detailed analysis of the DNA upstream and downstream of the actin genes did not lead to further information about the putative duplications.

		10	20	30	40	50	60	70	80	90	100	110
Act8	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp1	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp2	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp3	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp4	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp5	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp6	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp8	1	MESKVIVIH	GSHSLKIGLA	SESVPKTIPN	YIARKKKEKI	STPTITTET	PTIVSPVEST	IINNSNENN	DDIMKIDVEN	TVTPSEAVGT	TTELVKSTLP	MATTITITDEQ
Arp11	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
		120	130	140	150	160	170	180	190	200	210	220
Act8	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp1	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp2	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp3	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp4	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp5	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp6	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
Arp8	111	VKPTSSTSST	STTEEVEIKP	TESMDIDKPI	TDSTKNITSN	IKTEQPPPP	NIQTQTKIHI	PKKLEEEVEQ	SVKETTKLLP	PVDYIKVRQK	PTLNENNST	TFEIIKKKKK
Arp11	1	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----
		230	240	250	260	270	280	290	300	310	320	330
Act8	39	PRHTGVMVGM	GC-----	-----KDS--	YVGEAQSRR	G-----	-----LLTLRYE	TEHGIVTNWD	DMEKIWHHTF	YNELRVA--	EEHPVLLTE	AFLNKKANRE
Arp1	41	PKYKVMGLN	IE-----	-----NESTY	EVGDRNDWR	G-----	-----LLKLKHP	MDHGIVSNWS	DMERVVITYT	YDOLKIQ--	PSEHPVLLTD	VPNPERLRRE
Arp2	37	PILRSEEKVE	NV-----	-----EIKDI	MVGEAAKLR	S-----	-----MLQITYE	LENGIIRNWD	DITHVWDYAL	KEKLKVSDD	PTECKILLTE	PPMNEVANRC
Arp3	40	QSSKQKQTAA	SQKKGV----	-----EDLDE	FISDEAIAAS	K-----	-----TYDMANP	VKHSQIENWT	HMEQYWEHCV	EKVLRCE--	PEDYFLLTE	PPLNAPENRE
Arp4	40	VYKNENETVG	TGDSBMGEKD	D--SEPKRTY	YCCTNGITYR	RP-----	-----HMEIINE	ESDSLKNWD	AMEQIWNHTF	YERLHIN--	PIEHPILLAE	ETYNTRATRE
Arp5	40	YQCRAGFAND	ISPKLIFRSL	VGKVKSTSSP	IVNSNKEGD	IS-----	-----RLTIKSP	FDSNLLVEPP	QSESIFDYIS	EKEFGIEN--	EIENPVLLTE	ETSNETFCRK
Arp6	38	HTLKIATNNP	SQPHIIVPNQ	VGKVKNEKHQ	IMSEELINYN	DPS-----	-----EVRSRNE	MEKSYITNWG	LEKEIWDYAS	KRDDMKIK--	EQDNLLITE	APNSLDDLRK
Arp8	221	KTSSLSSLTN	VSTTPPPYVP	QPLNYNEIDY	CISDDAIAVS	RDKDKWFAYQ	PITVSTFNTG	IYHSVQSMED	DITQMKKPAI	QRYLNPSSD	LSSGCVYVV	TDNIDRKSLK
Arp11	37	LNISYLSGG	AS-----	-----FD	NFKKETSDN	ETI-----	-----TTTNIE	TKLPQSQLK	BSLFIFFKTI	ELNYLCK--	SDRKIIIVE	NIFRLRRE

		340	350	360	370	380	390	400	410	420	430	440
Act8	119	KMTQIMFET	NTPAMVVAIQ	AVLSLYASG	-----	-----	RTTGIVM	DSGDGVSHIV	PIVEGYALPH	AILRLDIAGR	DLTDYMMKIL	TERGYSFTTT
Arp1	122	RAAQLFETTY	NAPALYFSIP	AVLSLYASG	-----	-----	RTTGIVL	DSGDGVTHVV	PVFEFGALPH	AIISRIDAGR	DITEYLCHLL	RRSGYNEKTS
Arp2	120	KMTQIMFET	GFOAVVVAIQ	AVLTLYAQG	-----	-----	LLTGIVV	DSGDGVTHII	PVFEGYSLPH	LTRRLDVAGR	DVIRYLKLL	LLRGYAEENRT
Arp3	126	FTAEIMFET	NVEGLYIAVQ	AVLALAASWT	SKNAE	-----	KTLTGTVI	DSGDGVTHVI	PISEGVVIGS	SIKHLPIAGR	DISSYVOQIM	REREPNIPPA
Arp4	134	KVTEIMFET	HAPALEIAKN	AVLSSEFAS	-----	-----	KASSLV	DSGSGMTSVT	PVFEGEVVK	GIVKSNLAEN	RLTEEYRIL	SSKNITVNPI
Arp5	136	YMSSELLFECY	NIKSVVYGID	SLEFSYQQRD	QFKDG	-----	GKNSLII	GSSFNTHIY	NVQNYNVSHQ	QTKRINIGGG	ASTDYLRKLI	HLKYPKHKSY
Arp6	136	TMVEVVVEQY	KFKSLVLTTS	STEGLVHIKQ	QLLQYQQQHQ	QPPLDASNIS	LLKSPCHLVV	DCYSSTHII	PHCONTRUNY	AIKRFNIGCK	LLTNYLKEIV	SFR--YDMM
Arp8	331	QITITLLKEL	QFTSVLFFQ	SICSSSEGV	-----	-----	MATQSCVI	DLGHQKISHA	CVDEGYLLPN	TRLTGYGGE	QLTKLLEYLL	TGMDKSDSDI
Arp11	119	SMVSVLFEQY	RVFLVFINP	TASLIPTLR	-----	-----	TALLI	DSCFSETRVL	PIVEGVGLK	AYKSIISLSE	SLINQLKQYL	QSNITNGIII
		450	460	470	480	490	500	510	520	530	540	550
Act8	205	AEREIVRDIK	EKLAVVALD	-FEAEMQT	-----	-----	-----	-----	-----	-----	-----	-----
Arp1	208	AEKEVVRIIK	EKTCVVAHCP	QKEEELLE	-----	-----	-----	-----	-----	-----	-----	-----
Arp2	206	ADPETIRQIK	EKLCVVAID	-VQQEMKL	-----	-----	-----	-----	-----	-----	-----	-----
Arp3	219	ESLEIAKRVK	EQYSVVCPI	VKEFGKYD	-----	-----	-----	-----	-----	-----	-----	-----
Arp4	220	NLIKKEVEIKT	GEFGITPYNI	PTLTESYKRY	VT	-----	-----	-----	-----	-----	-----	-----
Arp5	228	FTQNYTNKIK	EEHCIVSQGQ	YIELKEFEN	DQLAKEKSVI	IQLPYQEIDF	EKLEERQQRK	IQNRKDLGAK	LRELADKKRL	EKKTELEDKL	ASLESILALK	TTNVEEFQQT
Arp6	244	HEIKLMNTIK	EKTCFISKCF	IFEIKRSQID	KLNSNLKIDY	VLPNYNDPNN	KTGYIKENLN	NNNNNND	-----	-----	-----	-----
Arp8	418	LTRQMVAKQI	HKYYEPFKSS	IYELVDSPFF	YLNVDNIKI	ENLDYYNDF	QKQRVGTFKV	KDIKHDKHMN	IYHFNADEVY	QVVGMSLFYP	NILSQFGGSS	VNYLIKRSRL
Arp11	204	DNNNQSITD	SMICNQLLND	KILNSIQLE	DIITR	-----	-----	-----	-----	-----	-----	-----
		560	570	580	590	600	610	620	630	640	650	660
Act8	230	-----	-----	-----	-----	-----	AASSS--AL	EKSYELPDGQ	-----	-----	-----	-----
Arp1	235	-----	-----	-----	-----	-----	PDSSSSKPV	QPOYTLPDGN	-----	-----	-----	-----
Arp2	231	-----	-----	-----	-----	-----	ASETT--VL	VENYTLPDGR	-----	-----	-----	-----
Arp3	246	-----	-----	-----	-----	-----	SEPKKIWT	INAQDSVTKK	PF	-----	-----	-----
Arp4	251	-----	-----	-----	LETRDL	KETSCRVS	DT	KDLGELINIA	QVPTELPGN	-----	-----	-----
Arp5	338	LKSKSYATEK	DLIRDIDDLK	DKLFGKKKES	EQVEDTEFP	LLFIADSELN	ADQLKPKKKQ	RQLKSMKDKR	LAQKRKRDEE	KEKEKEKEEE	RDRQEEESFL	KDPEHYLKDL
Arp6	310	-----	-----	-----	KNDKLVNI	EKDKDNNDIK	SKEEGEEIKL	NDIKKDDSTT	TTNTIKEDM	EQ	-----	-----
Arp8	528	ANTSESNIYV	EDQKHYYNHY	LSSYDHEDPF	DDHSILSFA	QNNTSRDNKD	GSNNNNIINN	NIINNIIINN	NNNNNNSSSS	SNNNNNNNS	GSNSNINSYN	NNNNNNNNNN
Arp11	238	-----	-----	-----	-----	CCFCFSK	NKNNNNIINN	NIINIRISKEY	S	-----	-----	-----

		670	680	690	700	710	720	730	740	750	760	770																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Act8	247	-----	-----	-----	-----	-----	-----	-----	-----	-----	VI	TIGNERERCP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp1	254	-----	-----	-----	-----	-----	-----	-----	-----	-----	VI	ELCAERERAP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp2	248	-----	-----	-----	-----	-----	-----	-----	-----	-----	VI	KVQGEREQAS																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp3	267	-----	-----	-----	-----	-----	-----	-----	-----	-----	SY	DVGYERELGP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp4	288	-----	-----	-----	-----	-----	-----	-----	-----	-----	LL	EIGSDREFQIP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp5	448	HSRKS	KILEK	REARQ	KQKQK	ANIVQ	RNSRL	RTIVN	PTNHG	NYGEK	GEEVE	DPEEA	EESRE	MAILD	KLLNK	FDPTS	ISSAI	VSHDD	QFPIG	EYHTA	EDFCV	SLG	VERTIKCP																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Arp6	361	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Arp8	638	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN	NNNNN																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																												
Arp11	266	-----	FN-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	NVTIVT	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS	NSTVP	STLNS

Amino acids

A G P S D E W Y H K R I L M V N Q T F C

Figure 10: The alignment shows the insertions and deletions (indels) of Act8 as compared to the Arps(Joseph *et al.*, 2008). The regions that align together show the clusters of similarity between the proteins and the perfectly aligned regions show the amino acids that form the key motifs of the actin protein. The amino acid color shading is shown in the bottom for reference.

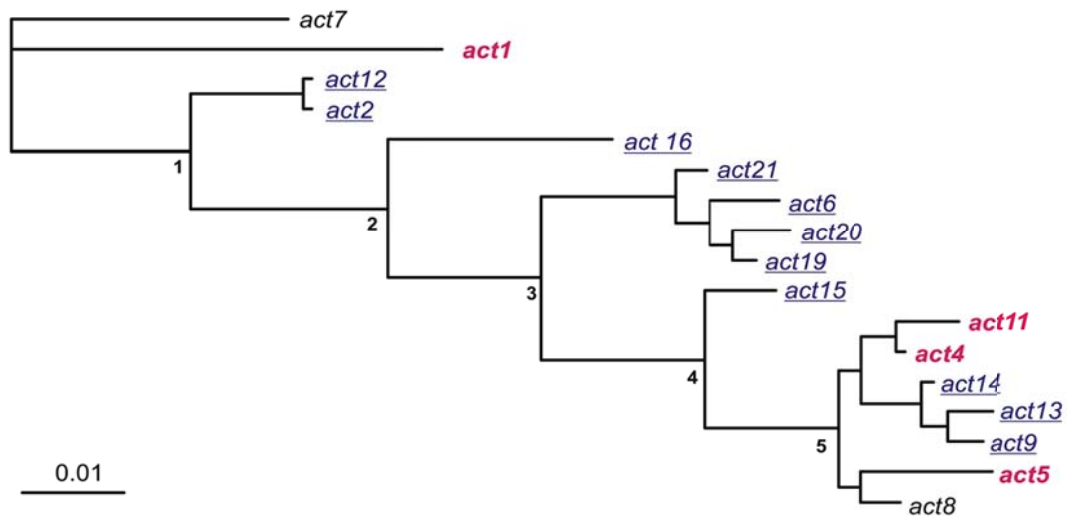


Figure 11: A tree analysis shows gross duplication events of the genes which code for identical actin proteins (Joseph *et al.*, 2008). Most of these genes are clustered on chromosome 2 (blue, underlined) and chromosome 5 (red, bold). An analysis of the corresponding genomic DNA implies five consecutive duplication events. The tree was generated with the program DNAML 3.5c.

3.4 Actin and Arps: structural features, models, and sequence logos

The crystal structure of the conventional actin was solved by Kabash *et al.* in 1990 (Kabsch *et al.*, 1990). Actin turned out to be a structural homolog of proteins like hexokinase, the Hsp70 family, other sugar kinases and prokaryotic cell cycle proteins such as MreB, FtsA and StbA (Doolittle and York, 2002). But with respect to the protein sequence, these proteins differ from actins although they follow the same molecular design (Kabsch and Holmes, 1995). It is believed that the products are the result of convergent evolution (Csete and Doyle, 2002). The conventional actins follow a typical actin fold, which comprises of four domains and five structural motifs.

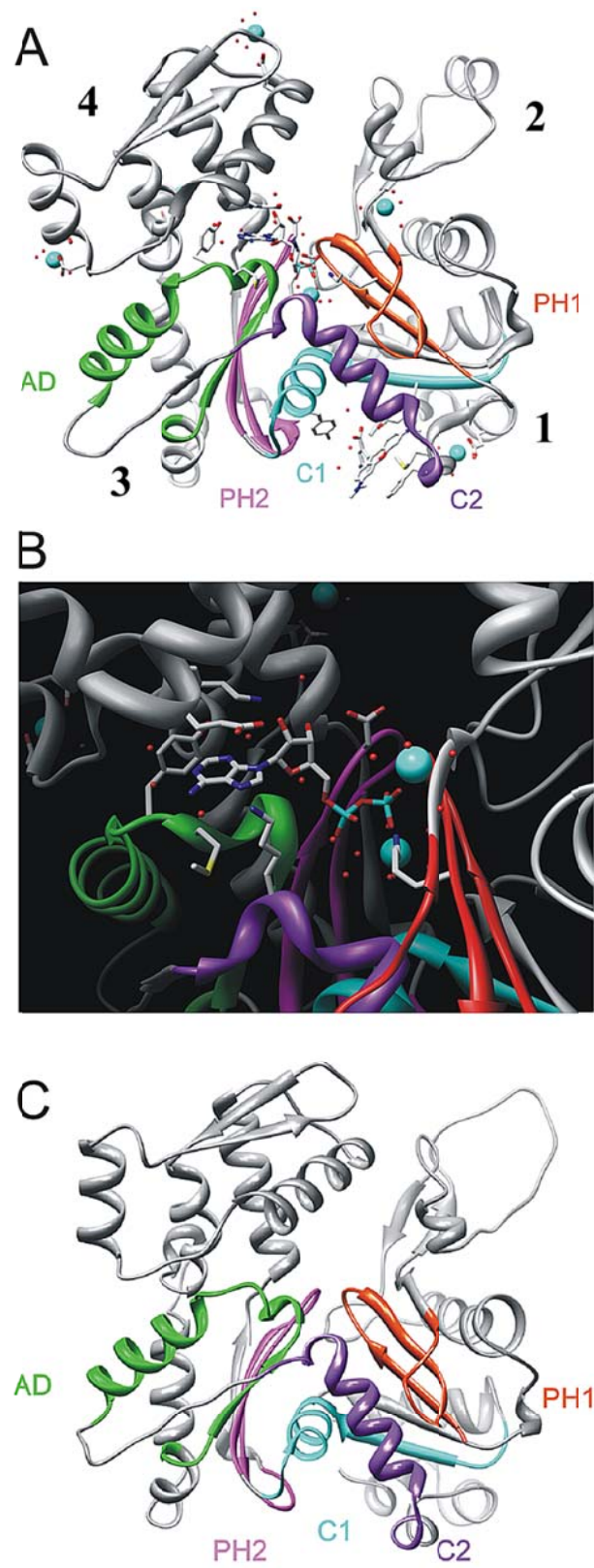


Figure 12: Actin crystal structure, ATP-binding motifs, and modeled Act8 structure (Joseph *et al.*, 2008). (A) The ribbon model of muscle actin (PDBID: 1J6Z) (Otterbein *et al.*, 2001) shows the wireframe of ADP, calcium ions (blue) and water molecules (red). The different domains have also been numerically represented from 1 through 4. (B) Five highly conserved sequences line the nucleotide binding pocket: the adenosine binding loop AD (green), the two phosphate binding loops PH1 and PH2 (red and purple, respectively) and the subdomain connecting motifs C1 and C2 (blue and violet, respectively). These sequences are highly conserved signatures in actins and actin-related proteins. (C) Model of Act8, the representative of the 17 identical actins in *D. discoideum*.

3.4.1 Domains

The core of actin is formed from two structurally similar globular domains (domain I and domain II) connected by a flexible hinge. The interface between the two domains forms an ATP-binding pocket, which plays a primary role in the dynamics of actin. Each domain is composed of two subdomains 1 (Ia), 2 (Ib) and 3 (IIa), 4 (IIb). The subdomains 1 and 3 define the 'barbed end', where capping proteins bind actin, as opposed to the 'pointed end', composed of subdomains 2 and 4 (Carlier, 1990). The major interactions between actins and the binding partners is mediated by a hydrophobic cleft in between the two largest subdomains 1 and 3. Evolution: The structural similarity between the two domains of the G-actin suggests that they were formed by the duplication of one ancestral domain (Bork *et al.*, 1992).

3.4.2 Motifs

The five motifs (Bork *et al.*, 1992) are the adeno-motif (AD), two phosphate binding motifs (PH1 & PH2) and two connection motifs (C1 & C2). These motifs were selected for analysis. Based on actin these motifs were identified for all the members of the actinome. Using 'sequence logos' we identified the most conserved residues found in these motifs (Fig. 13,14). The sequence logos (Crooks *et al.*, 2004) of the members of the actinome show the key residues, which are obviously of structural, functional and evolutionary significance (Bork *et al.*, 1992).

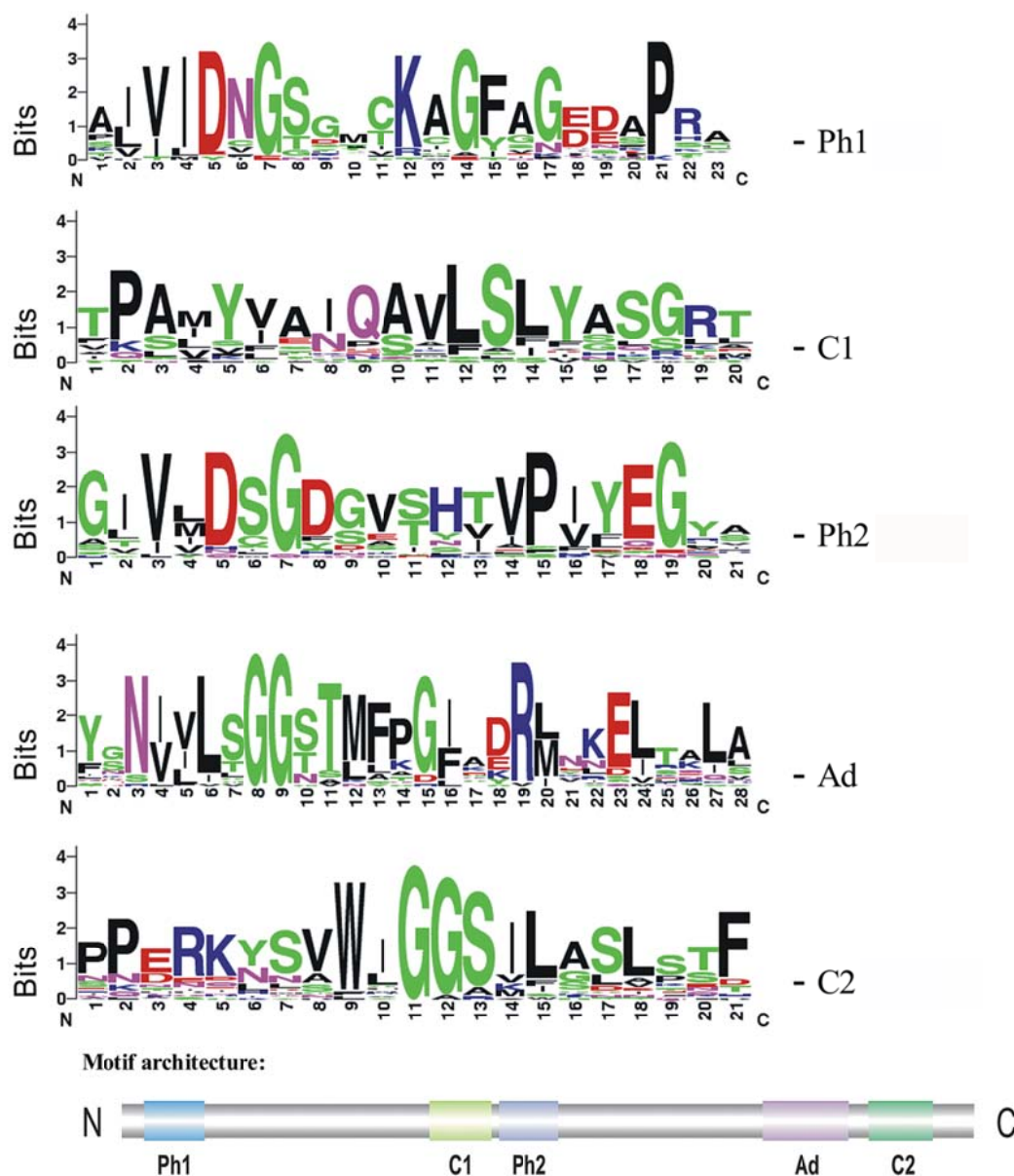


Figure 13: Motif logos from *D. discoideum* actins and Arps (Joseph *et al.*, 2008). The five structural motifs of all members of the *D. discoideum* actinome are summarized as sequence logos which reflect the structural, functional and evolutionary significance of specific amino acids at a particular position. It is remarkable that there is a surprisingly pronounced variability of conservation and only three glycine residues are seemingly indispensable in all 41 members of the *D. discoideum* actinome (# 8 and #9 in the adeno domain, # 11 in C2).

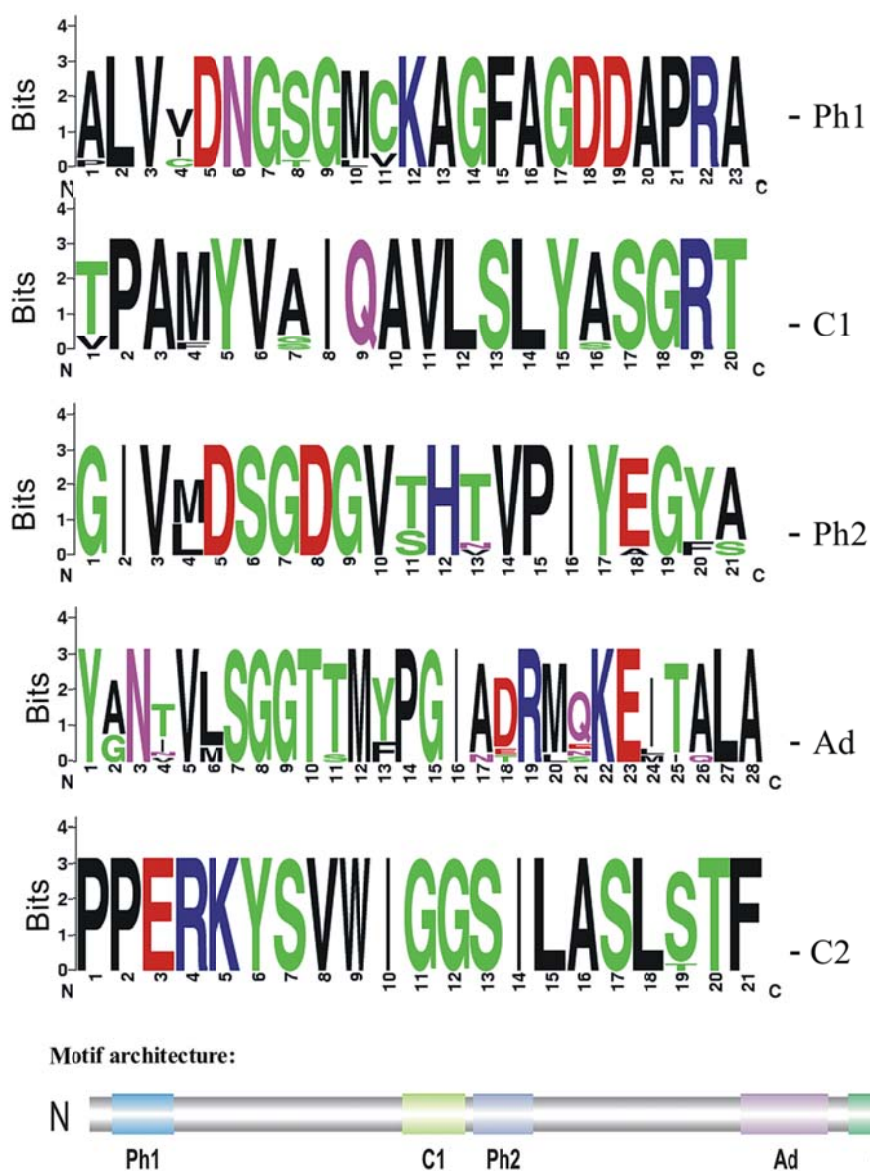


Figure 14: Motif logos from Act8 homologs across species (Joseph *et al.*, 2008). The five structural actin motifs are summarized as sequence logos which reflect the structural, functional and evolutionary significance of specific amino acids at a particular position. In difference to the variability of the motif profiles in *D. discoideum* only (Fig. 13), the conventional actins are extremely conserved from lower to higher eukaryotes.

The homologies in conventional actins throughout evolution are extremely high(Erickson, 2007). The following actins have been compared: *Arabidopsis thaliana*, Q541W9_ARATH - *Caenorhabditis elegans*, ACT2_CAEEL - *D. discoideum*, Act8, DDB0216213 - *Drosophila melanogaster*, ACT1_DROME(Act5C) - *Entamoeba histolytica*, Q51B76_ENTHI - *Gallus gallus*, ACTB_CHICK (beta-actin) - *Homo sapiens*, ACTG_HUMAN (gamma-actin) - *Mus musculus*, ACTG_MOUSE (gamma-actin) - *Saccharomyces cerevisiae*, ACT_YEAST - *Takifugu rubripes*, ACT1_FUGRU, (beta-actin1) - *Xenopus laevis*, ACTB_XENLA (beta-actin).

All actin isoforms when modeled onto the known structure of actin, exhibit similar structural features to actin, whereas actin-related proteins differ drastically in several cases (Fig. 15).

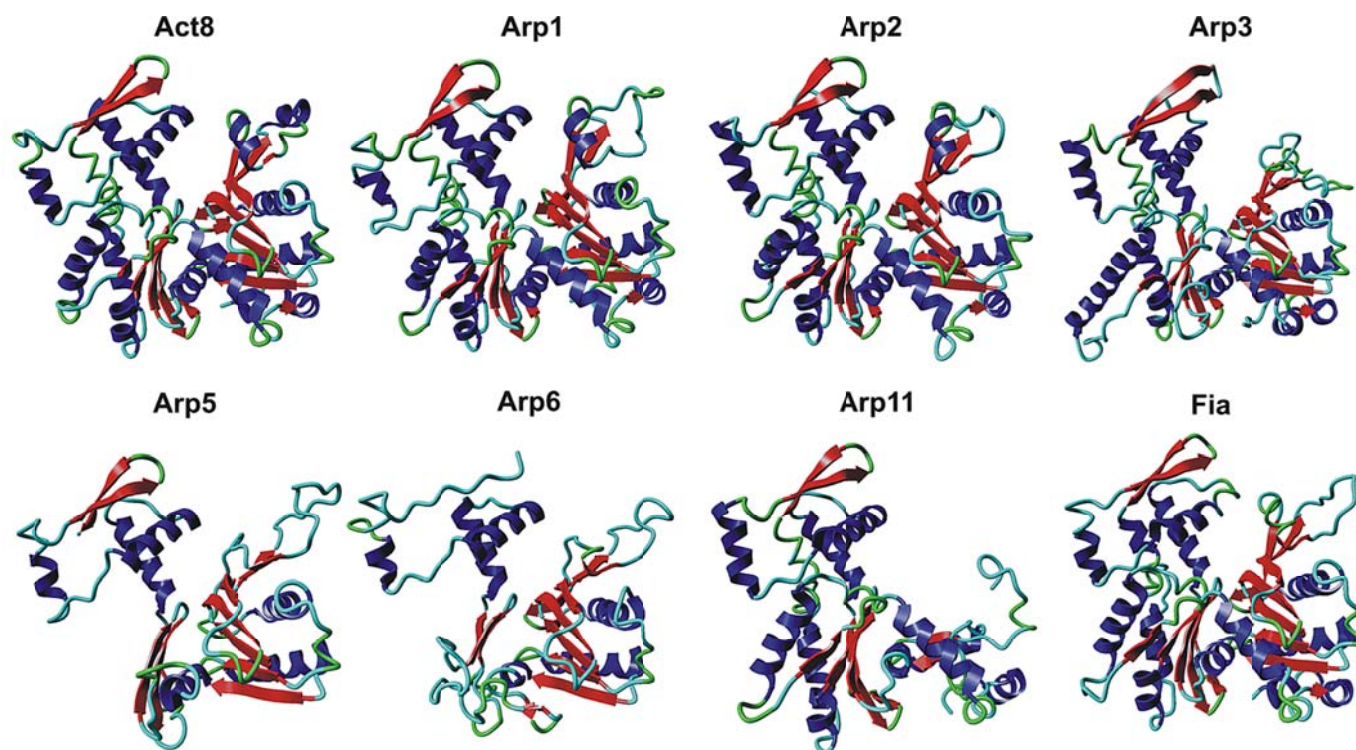


Figure 15: Structural homology among the *D. discoideum* Arps (Joseph *et al.*, 2008). The sequences were modeled in comparison to the Act8-type actin (upper left). Whereas Arps 1, 2 and 3 show high similarity to the three-dimensional structure of conventional actin, Arp 5, 6 and 11 are clearly different. Attempts to model Arps 4 and 8 failed due to low structural homology and the absence of available templates. Filactin (Fia, lower right) shows high structural homology in the actin domain. Modeling was done using the swiss model server, graphics were generated with the YASARA molecular visualization software.

3.5 Localization signals and sorting motifs

3.5.1 Nuclear localization signal (NLS)

NLS(Cokol *et al.*, 2000) is a short stretch of amino acids that mediates the transport of nuclear proteins into the nucleus. It is abundant in positively charged residues. A NLS could contain monopartite and bipartite motifs. The members of the actinome were processed with the PredictNLS server looking for NLS motifs. NLS motifs were detected only in two sequences Arp5 (KKKQR QLKSM KDGRL AQKRKR) and Arp8 (RKKKEK) as putative nuclear proteins. This correlates well with own and other's data and reports that identified these two Arps as members of chromatin remodeling complexes(Shen *et al.*, 2003).

3.5.2 Nuclear export signal (NES)

In contrast to the nuclear localization signal, the NES motif(la Cour *et al.*, 2004) is responsible for nuclear export. NLS has been shown as a collection of patterns that can be identified *in silico*, but identifying a NES is significantly more complicated and consensus patterns unfortunately do not suffice. We used the NetNES server to predict the nuclear export signals in the members of the actinome. We also used the information collected from the work of E. Nishida and coworkers(Wada *et al.*, 1998) as a base to map other NES in the members of the actinome. The results are shown in Table 4.

Table 4: The putative nuclear export signals are shown (Joseph *et al.*, 2008). The signals according to Wada *et al.* (1998) have been identified manually. Sequences in brackets had too weak characteristics to define them as explicit nuclear export signals.

SI. NO	Proteins	Signals predicted using the NetNES 1.1 server	Signals predicted based on Wada <i>et al.</i> 1998 (Wada <i>et al.</i> , 1998) (top: NES1, bottom: NES2)
1	Act3	176-ILR LDLAGRD LDY-189 219-YVA LD FE-225	171-SLP HA ILR LDLAG -183 212-D I KE K L A Y V ALD-223
2	Act10	176-ILR LDLAGRD LDY-189 219-YVA LD FE-225	171-ALP HA ILR LDLAG -183 212-D I KE K L A Y V ALD-223
3	Act8*	176-ILR LDLAGRD LDY-189 219-YVA LD FE-225	171-ALP HA ILR LDLAG -183 212-D I KE K L A Y V ALD-223
4	Act17	216-KLS YIT LDFQ-225	171-SIN HA IS R L DLAG -183 212-D I KE K L S Y I T L D-223
5	Act18	178-RLD LAGRD LSDY-189 219-YVA LD FD-225	171-ALP HA ILR LDLAG -183 212-D I KE K L S Y V ALD-223
6	Act22	176-ILR LDLAGRD LDY-189 219-YVA LD FE-225	171-ALP HA ILR LDLAG -183 212-D I KE K L A Y V ALD-223
7	Act23	160-ILR LDLAGRD LDY-173 203-YVA LD FE-209	155-ALP HA ILR LDLAG -163 196-D I KE K L A Y V ALD-207
8	Act24	178-RLH LAGG -184 220-YVA LD FE-226	171-TFPL SITR L HLAG -183 213-D I KE K L A Y V ALD-224
9	Act25	190-IL S L DLAGR -199 234-YVA LD FE-240	186-ILRQ AIL S L D LAG -198 227-G I KE K L S Y V ALD-238

10	Act26	218-YVA I NFN-224	170-VLQHS I IRLN F AG-182 211-S I KEK L AY V AIN-222
11	Act27	NIL	160-ALPHATSS L GFAG-172 197-D I KEK L SY V SSD-208
12	Act28	NIL	162-ALPLATSS L GFAG-174 199-D I KEK L SY V SSD-210
13	Act29	NIL	177-S I PNS I NQ L EIAG-189 219-D I KEK F GF I SLN-230
14	Filactin	20-RLV L HKN-25	739-S I PHA I KR I D I GG-751 780-D I KEK T S F V SQD-791
15	Arp1	287-DMS I RKS-293	174-ALPHA I SR I D I AG-186 215- I KEK T CY V AHD-226
16	Arp2	NIL	172-S I PHL T RR L D V AG-184 213-Q I KEK L CY V AYD-224
17	Arp3	57-DLD F FIG-63	185-V I GSS I KH I P I AG-197 226-R V KEQYSY V CPD-237
18	Arp4	NIL	186-V V KNG I VKS N LAG-198 (SYKRYVTLET I R)
19	Arp5	323-ILAL L KTT-329	284-RQRKIQNRKD L GA-296 (T LKSKSYATEKD)
20	Arp6	425-KER L E L E L RLKL-435	212-RLNYA I KRFN I GG-224 251-T I KEK T CF I SKD-262
21	Arp8	334-TLL L KEL-340	384-LLPNTRL T LG Y GG-396 (DFSDNIKIENLD)
22	Arp11	80-KES L FIF-86	170-G I LKAYKS I SLGS-182 219-Q L LNDKILNS I Q-230

23	Act 31 (355aa)	NIL	164-PVTDAAVVTLDFFGG-176 205-QIKEKHSFIELD-216
24	Act32 (392aa)	227-YVSLNLYN-233	179-IITKAMSHLPLGG-191 220-DIKEKLGYSVSLN-231
25	Act33 (414aa)	360-RLKIELG-366	190-LLKEGIVRQEFEGG-202 NIL

3.5.3 Signal peptide prediction (SignalP)

The SignalP 3.0 server(Nielsen *et al.*, 1997)predicts the presence and location of signal peptide cleavage sites in amino acid sequences from different organisms: gram-positive and gram-negative prokaryotes, and eukaryotes. This method incorporates a prediction of cleavage sites and a signal peptide/non-signal peptide prediction based on a combination of several artificial neural networks and hidden markov models. The *D. discoideum* actinome was posted to the SignalP server, which as expected did not find any members of the actinome to have a signal peptide.

3.5.4 Trans membrane hidden markov model (TMHMM)

The TMHMM(Krogh *et al.*, 2001) is a server that identifies the transmembrane helices in proteins. It looks for membrane spanning regions and the orientation of the protein. The members of the actinome were analyzed in this server for the presence of transmembrane helices. None of them were identified to have a transmembrane helix, which implies that all of the members of the actinome, also the structurally rather distant actin-related proteins, are soluble proteins and not membrane proteins.

3.5.5 Proline(P) Glutamic acid(G) Serine(S) Threonine(T) rich regions (PEST motifs)

PEST motifs(Rogers *et al.*, 1986; Rechsteiner and Rogers, 1996) are rich in the amino acids P, E, S and T, which reduce the half-lives of proteins dramatically and target them for proteolytic degradation. Table 5 contains a list of proteins from the actinome that harbor PEST motifs. Among them are the centrosome-specific Arp1 (contractin), the

putative subunits from nuclear complexes (Arp5, Arp8), and Filactin. The latter is a conventional actin with a long N-terminal extension. As observed in our laboratory Filactin is notoriously difficult to purify due to its biochemical instability.

Table 5: Members of the actinome that harbor PEST motifs (Joseph *et al.*, 2008). The corresponding amino acid sequences are shown on the right.

SLNO	Proteins	PEST Motifs
1	Filactin(224-251)	KESTATIDQFPSPPTSNISTTSTTTTTT
2	Arp1(229-241)	KEEELLEPDSSSS
3	Arp5(491-505)	KGEEVEDPEEAEES
4	Arp8 (75-95) (111-128)	KIDVENTVTPSEAVGTTTEDV KPTSSTSSTSTTEEVEI
5	Arp11(54-71)	KEITSDNETITTTNQIPT

3.6 Actin homologs across species

BLAST search was used to look for homologs of *D. discoideum* Act8 across species. The best scores are shown in Table 6. While the lower organisms have just actin, the actin isoforms (Rubenstein, 1990) beta and gamma become best hits in higher organisms. *X. laevis* and *T. rubripes* have beta-actins as their best hits. *G. gallus*, *M. musculus*, and *H. sapiens* have gamma actin as their best hits.

Table 6: The *D. discoideum* actinome as compared to all at that time available genomes(Joseph *et al.*, 2008). The blast scores, protein ID, Uniprot ID are all shown. Especially interesting are the best hits of the isoforms that rank first in the blast results as they are closer to *D. discoideum*.

SI. NO	Organism	Protein Identification	Blast Scores (Bits)	E value	Best hit (isoform)	Uniprot Identification
1	<i>E. histolytica</i> (Loftus <i>et al.</i> , 2005)	Eh_ACT	695	0.0	actin	P11426
2	<i>S. cerevisiae</i> (Hirschman <i>et al.</i> , 2006)	Sc_ACT1	686	0.0	actin	P60010
3	<i>A. thaliana</i> (McKinney and Meagher, 1998)	At_ACT11	712	0.0	actin	P53496
4	<i>C. elegans</i> (Stein <i>et al.</i> , 2003)	Ce_ACT-2	724	0.0	actin	P10984
5	<i>D. melanogaster</i> (Celniker and Rubin, 2003)	Dm_ACT5C	724	0.0	actin	P10987

6	<i>X. laevis</i>	Xl_ACTB	725	0.0	beta-actin	O93400
7	<i>T. rubripes</i>	Tr_ACTB-A	723	0.0	beta-actin	P68142
8	<i>G. gallus</i> (Consortium, 2004)	GG_ACTG1 Gg_ACTB	724 724	0.0 0.0	gamma- actin	Q5ZMQ P60706
9	<i>M. musculus</i> (Pena-Castillo <i>et al.</i> , 2008)	Mm_ACTG1 Mm_ACTB	724 724	0.0 0.0	gamma- actin	P63260 P60710
10	<i>H. sapiens</i> (Strausberg <i>et al.</i> , 2002)	Hs_ACTG1 Hs_ACTG2 Hs_ACTB	724 724 724	0.0 0.0 0.0	gamma- actin	P63261 P63267 P60709

3.7 Actin Related Proteins (Arps)

The Arp(Frankel and Mooseker, 1996)families were first identified(Poch and Winsor, 1997) in *S. cerevisiae* and classified as Arp1 to Arp10 based on the descending order of their sequence similarity to actin(Boyer and Peterson, 2000). The additional Arp11 was then identified in the human genome (Frankel and Mooseker, 1996). Apart from this, there is a family of orphan Arps. The Arps in *D. discoideum* were identified and a comparative study was performed with the sequenced genomes of other model organisms. Most organisms contain single copy Arp genes, with the exception of *E. histolytica*, which has two genes that encode an Arp5 (NCBI-Database). The *X. laevis* genome contains two Arp2 isoforms (NCBI-Database) and a few yeast species have two copies of Arp4(Muller *et al.*, 2005). With just four Arps (Arp2, 3, 4, 5) the parasite *E. histolytica* has only a small repertoire of actin related proteins. However, the host cell might compensate for the absence of other ubiquitous Arps. *T. rubripes* has the most compact genome, which contains only two Arp genes encoding Arp3 and Arp5, so far the smallest number of actin-related proteins in a sequenced genome (Table 6).

Table 7: The presence of the Arps in different organisms (*assembly) that represent various taxonomical niches(Joseph *et al.*, 2008)

Sl.NO	Organisms	Arp1	Arp2	Arp3	Arp4	Arp5	Arp6	Arp7	Arp8	Arp9	Arp10	Arp11
1	<i>A. thaliana</i>	-	+	+	+	+	+	-	+	-	-	-
2	<i>E. histolytica</i>	-	+	+	+	+	-	-	-	-	-	-
3	<i>D. discoideum</i>	+	+	+	+	+	+	-	+	-	-	+
4	<i>S. cerevisiae</i>	+	+	+	+	+	+	+	+	+	+	-
5	<i>D. melanogaster</i>	+	+	+	+	+	+	-	+	-	-	+
6	<i>C. elegans</i>	+	+	+	+	-	+	-	-	-	-	+
7	* <i>X. laevis</i>	+	+	+	+	+	-	-	+	-	-	+
8	* <i>T. rubripes</i>	-	-	+	-	+	-	-	-	-	-	-
9	* <i>G. gallus</i>	+	+	+	+	+	+	-	+	-	-	+
10	<i>M. musculus</i>	+	+	+	+	+	+	-	+	-	-	+
11	<i>H. sapiens</i>	+	+	+	+	+	+	-	+	-	-	+

3.8D. *discoideum* actin in comparison to actins from other organisms

Figure 16 shows the relationship of *D. discoideum* conventional actin with actin representatives from other organisms. The sequences used are all Arp8 homologs of *D. discoideum* in the respective organisms. Due to incomplete taxon sampling, this figure reflects only a rough estimation.

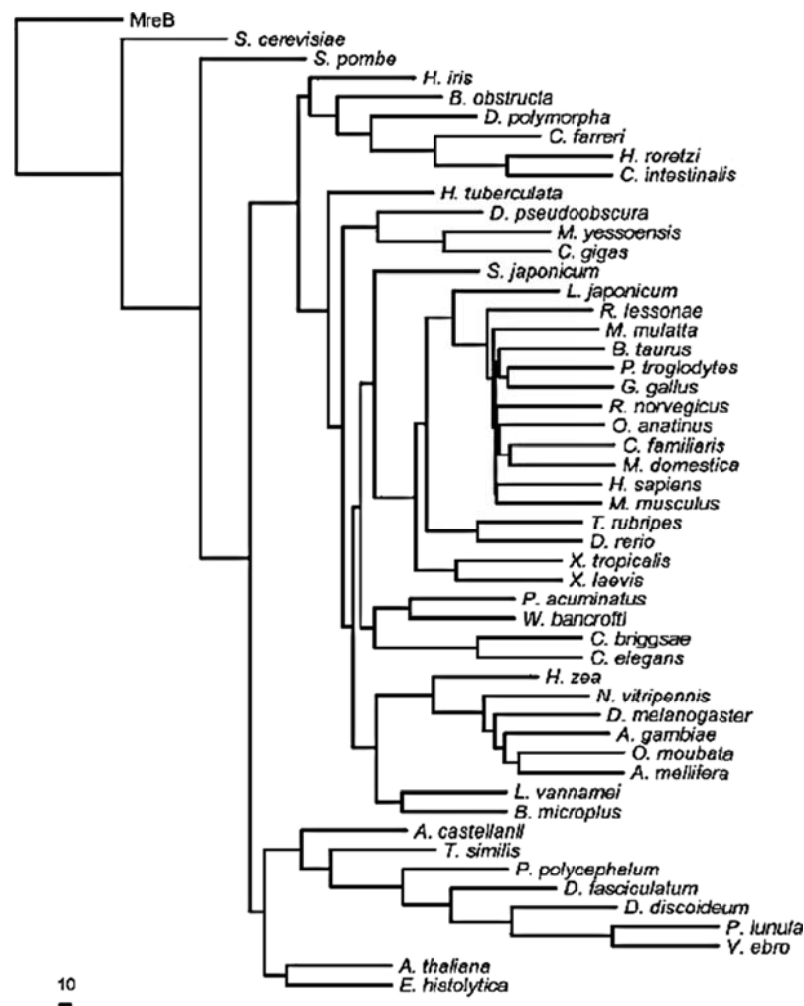


Figure 16: A wider view of actin homologs from all available organisms. The analysis was done with the conventional actin from *D. discoideum* (Act8), the bacterial actin-like protein MreB was used as an outgroup, and the overall organization of the tree shows clearly the extremely conserved nature of actin throughout evolution. The following organisms have been included in this analysis from top to bottom: *Drosophila*

pseudoobscura, *Mizuhopecten yessoensis*, *Crassostrea gigas*, *Schistosoma japonicum*, *Lethenteron japonicum*, *Rana lessonae*, *Macaca mulatta*, *Bos Taurus*, *Pan troglodytes*, *Gallus gallus*, *Rattus norvegicus*, *Ornithorhynchus anatinus*, *Canis familiaris*, *Musca domestica*, *Homo sapiens*, *Mus musculus*, *Takifugu rubripes*, *Danio rerio*, *Xenopus tropicalis*, *Xenopus laevis*, *Phacus acuminatus*, *Wuchereria bancrofti*, *Caenorhabditis briggsae*, *Caenorhabditis elegans*, *Helicoverpa zea*, *Nasonia vitripennis*, *Drosophila melanogaster*, *Apis mellifera*, *Litopenaeus vannamei*, *Boophilus microplus*, *Acanthamoeba castellanii*, *Trachypenaesus similis*, *Physarum polycephalum*, *Dictyostelium fasciculatum*, *D. discoideum*, *Pyrocystis lunula*, *Vannella ebro*, *Arabidopsis thaliana*, *Entamoeba histolytica*.

3.9 Actin mutations that lead to human diseases, in relation to distinct actin variations in the *D. discoideum* actinome

Several actin mutations have been found in humans (Costa *et al.*, 2004). Two of these naturally occurring mutations have been studied in detail. They are:

Alpha 1, skeletal muscle, ActA1; which is relatively frequent.

Actin, alpha cardiac, ActC1; which is relatively rare and only two of them have been reported.

ActB mutation has been reported in a single instance and in a single case.

In *D. discoideum* the members of the actinome were searched for harboring these residues. The results are shown in Table 8. The relative number of ESTs (Expressed Sequence Tags) present for each of the protein is also listed to show the approximate level of expression of these proteins (none=0, X=1-5, XX=6-15, XXX= >15). The two ActC mutations (Arg312His and Glu361Gly) have not been identified among the members of the actinome. Critical mutants highlight the crucial residues and vital positions in the actin molecule. Research in disease mutants of actin can yield valuable information on the molecular function.

Table 8: ActA1 mutations(Costa *et al.*, 2004), which are found in members of the *D. discoideum* actinome. These mutations are clinically significant in humans as in most cases they cause diseases. There could be other mutations in human actins at positions other than those mentioned in the table below, and amino acid replacements other than those in the table; these mutations do not cause any diseases.

Sl.NO	Mutation	Found in <i>D. discoideum</i> actin sequences	rel. Number of ESTs
1	G15R	Arp11	0
2	H40Y	Act3, Act17, Act18, Act33, Arp1	X,XX,X,0,XXX
3	I64N	Act33	0
4	L94P	-	
5	N115S	Arp8(alignment ambiguous)	X
6	M132V	Arp2	XXX
7	I136M	-	
8	V163L	Arp8,11 (alignmet ambiguous)	X,0
9	G182D	-	
10	R183G	Act24	XXX
11	R183C	Act27, Act28	0,X
12	G259V (Error)	There is an E instead of G at that position. There is a G at 251 but no V replacements are found.	
13	Q263L	Act27, Act28	0,X
14	G268C	-	
15	G268R	-	
16	N280K	Act26, Arp8	0,X
17	D286G	-	
18	I357L	-	
19	V370F	-	

Initially the whole actinome (Joseph *et al.*, 2008) was studied through bioinformatic means only (*in silico*). Later selected proteins were chosen for more detailed experimental studies. The experiments started with PCR and molecular cloning. Biochemical experiments like expression and purification of the proteins, as well as antibody generation followed. At the cell biology level knockout and GFP-fusion constructs were transformed into *D. discoideum*. Microscopy was done with the purified antibodies and strains which overexpressed GFP fusion constructs (data not shown).

3.10 Actin related protein 5 (Arp5)

Arp5 is a nuclear Arp (Shen *et al.*, 2003), which in yeast is associated with chromatin (Grava *et al.*, 2000). It is part of the INO80 complex (Jin *et al.*, 2005). Very little work has been done on this Arp in other organisms and it has been never worked on in *D. discoideum*.

As a first step antibodies were generated against a selected region of Arp5, spanning the amino acids #408 - #550. The Arp5 peptide was expressed as a Glutathione-S-transferase fusion protein and used for antibody production in rabbits. Since the members of the *D. discoideum* actinome are nearly identical in many regions, it was important to exclude cross-reactions of the antibodies with other members of the actinome. Therefore, care was taken to identify a region specific to Arp5. The members of the actinome were aligned as shown in the Figure 17 and Arp5 analyzed for unique protein stretches.

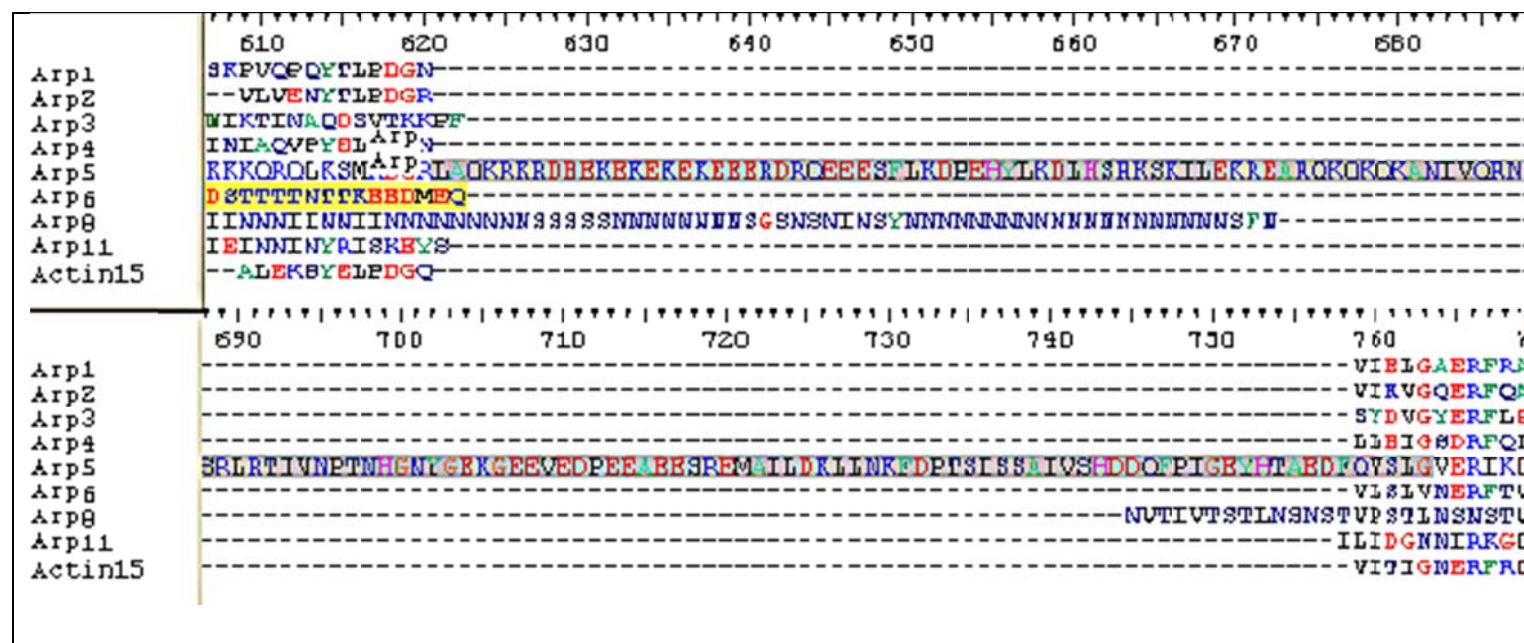


Figure 17: Distinct regions of Arp5 for an increased specificity of the antibodies generated. The piece for Arp5 selected for expression is highlighted in grey among the multiple alignment of the different Arps. This region includes the Arp5 polypeptide from 408-450. Unfortunately, these numbers do not correspond to the numbers in the figure, which

represent the total alignment from different Arp proteins. The Arp6 is the only other Arp protein which has been expressed in this study, so as to represent the relative position of Arp6, it has been highlighted in yellow.

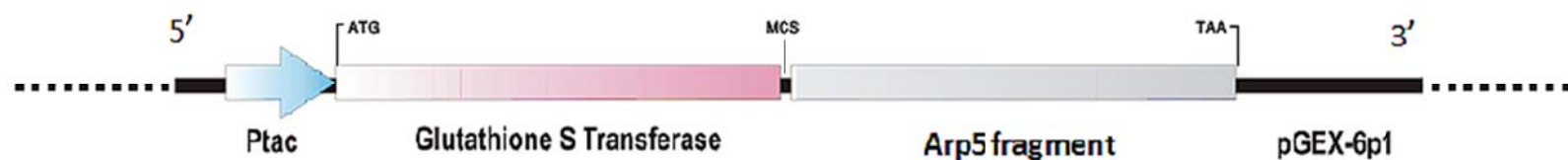


Figure 18: The scheme of the construct of Arp5 fused to GST on a pGEX-6p1 vector. The Arp5 fragment is cloned in frame with the Glutathione S transferase at the multiple cloning site (MCS) and TAA is added to the end of the Arp5 fragment.

The Arp5 PCR fragment was cloned with the restriction sites EcoRI & SalI. The insert was sequenced along with the GST gene and no discrepancies were found.

Several cloned transformants were induced with IPTG after the OD₆₀₀ reached a value around 0.6. Clones #1 and #10 showed additional bands at 43 kDa as compared to their respective uninduced lanes. The Arp5 fragment has an expected mass of 17 kDa, GST alone is 26 kDa.

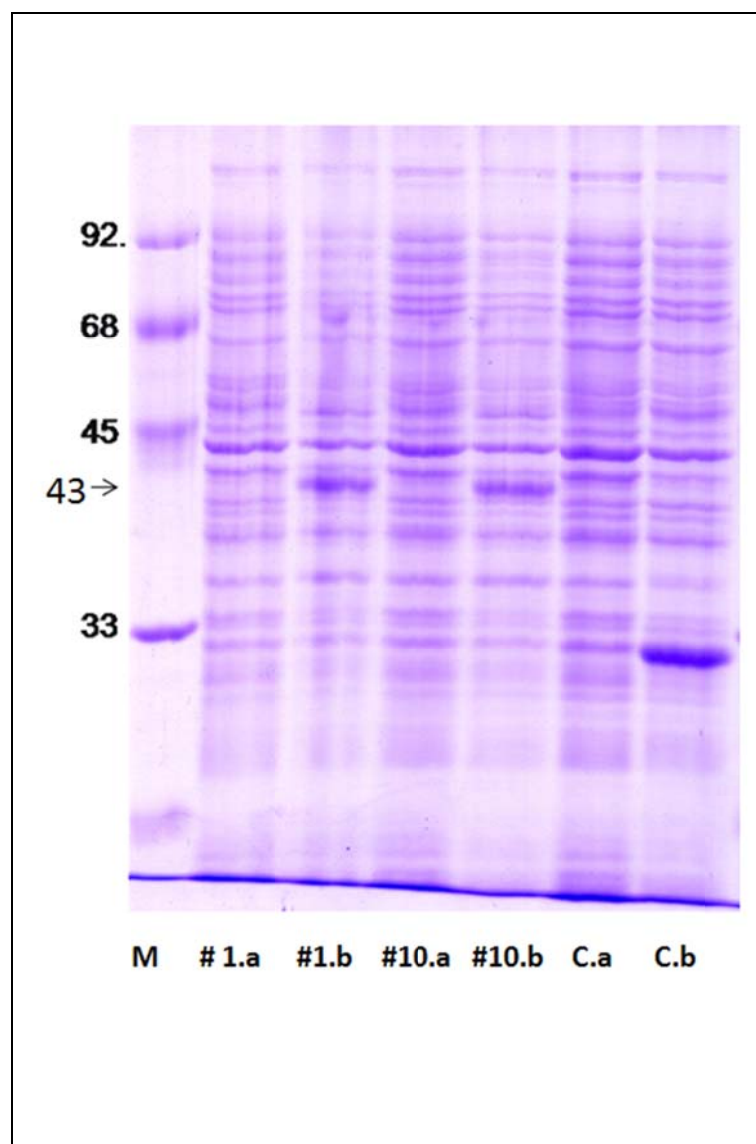


Figure 19: The expression pattern of the bacterial clones #1 and #10 have been depicted at the figure above. For example #1.a is a uninduced clone number 1, while #1.b is an induced clone number 1. This is also true with the controls where C.a is a bacterial lysate before and C.b is a bacterial lysate after induction of GST. M stands for the molecular weight markers (ladder). The bands at about 43 kDa (arrow) consist of the expressed GST-Arp5 fusion protein.

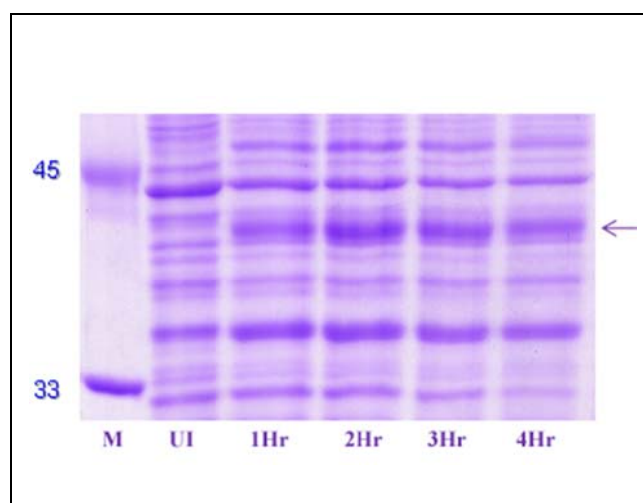


Figure 20: Expression of fusion protein in clone #10 over 4h of induction. 45 and 33 show the molecular weights of the marker proteins in the first lane (M). The arrow points at the band of interest. In comparison to the uninduced sample (UI), expression is visible already after 1h (1Hr) and seems to peak after about 2h (2Hr) of induction.

The Arp5-GST fusion peptide from clone #10 was purified following standard procedures and sent for production of rabbit antibodies to the company Pineda (Berlin, Germany). Antisera were tested over periods of 60, 90 and 150 days after the first immunization. The last serum was found to be the best of all. In Figure 21 the Coomassie blue stained SDS polyacrylamide gel and the Western blot are shown side by side. Lane 2 contains a whole homogenate from AX2 cells, lanes 3 and 4 contain the protein fraction which was used for immunization at very high dilutions. The protein used for immunization clearly shows a reaction with the antiserum (arrow), the higher band might be a running artifact. The band in lane 2 suggests a reaction with endogenous Arp5 that has a predicted molecular mass of 79 kDa.

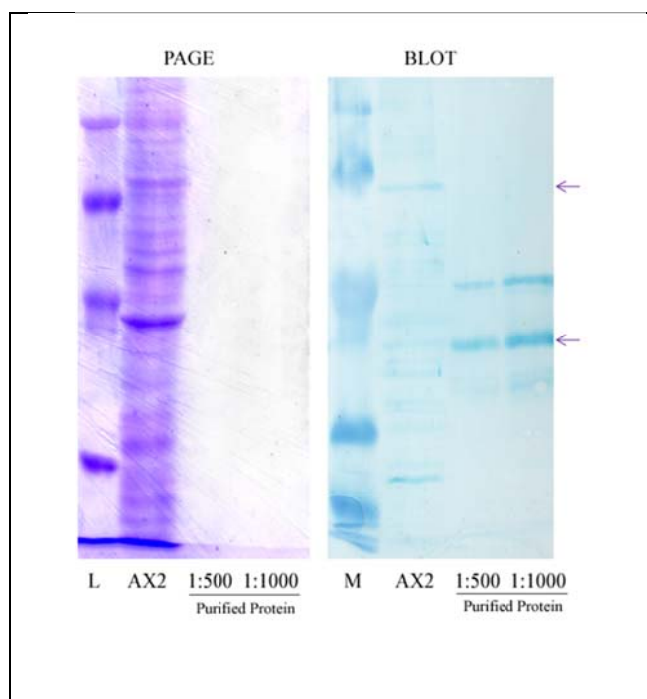


Figure 21: Test of the polyclonal antiserum. The left panel shows a Coomassie blue stained SDS gel (12% acrylamide) with a whole cell homogenate (AX2 wild type) and two different dilutions of a protein fraction with purified recombinant protein. The corresponding Western blot shows that the recombinant protein despite its low concentration is recognized by the antibodies (lower arrow), and that a band in the homogenate (upper arrow) might represent the endogenous Arp5. Lane L contains the conventional marker sample with molecular weights of 92 kDa, 68 kDa, 45 kDa and 33 kDa from top to bottom, for the blot a prestained marker (M) with the molecular weight of 122 kDa, 76 kDa, 47 kDa, 23 kDa and 20 kDa from top to bottom.

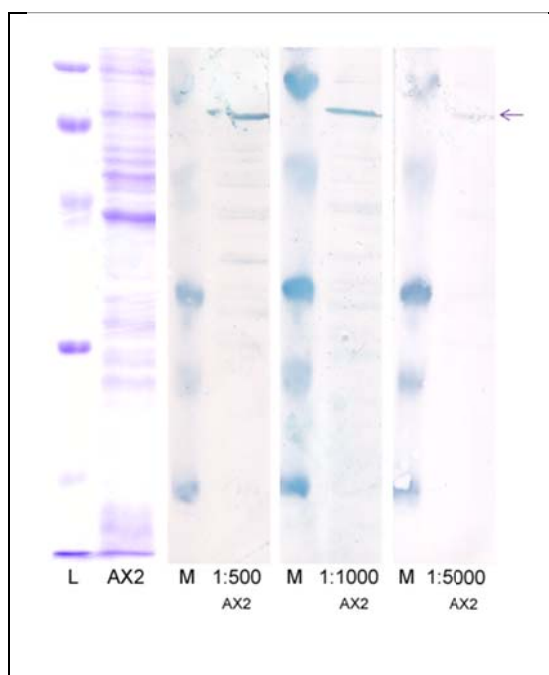


Figure 22: Detection of Arp5 with affinity purified polyclonal antibodies. The two lanes at the left show the molecular weight markers and the whole cell lysate (2×10^5 cells/lane) after staining the gel with Coomassie blue. The three panels at the right depict the Western blots and the crossreaction with endogenous Arp5 (arrow) after separation of lysates from the same number of cells but after incubation with three dilutions of affinity purified antibody. Lane L contains the conventional marker sample with molecular weights of 92 kDa, 68 kDa, 45 kDa and 33 kDa from top to bottom; the blot has prestained markers indicated as M, with the molecular weights of 76 kDa, 47 kDa, 23 kDa and 20 kDa from top to bottom.

The affinity purified antibody was used also in immunofluorescence studies. Fig. 23 shows a typical experiment in which attached cells have been fixed with methanol and stained with phalloidin for detecting the distribution of actin, with DAPI as a counter stain for the nucleus and with affinity purified Arp5 antibody. Arp5 was found almost exclusively in the nucleus which confirmed the data from other groups who described the nuclear localization of Arp5 in yeast (Meagher *et al.*, 2009).

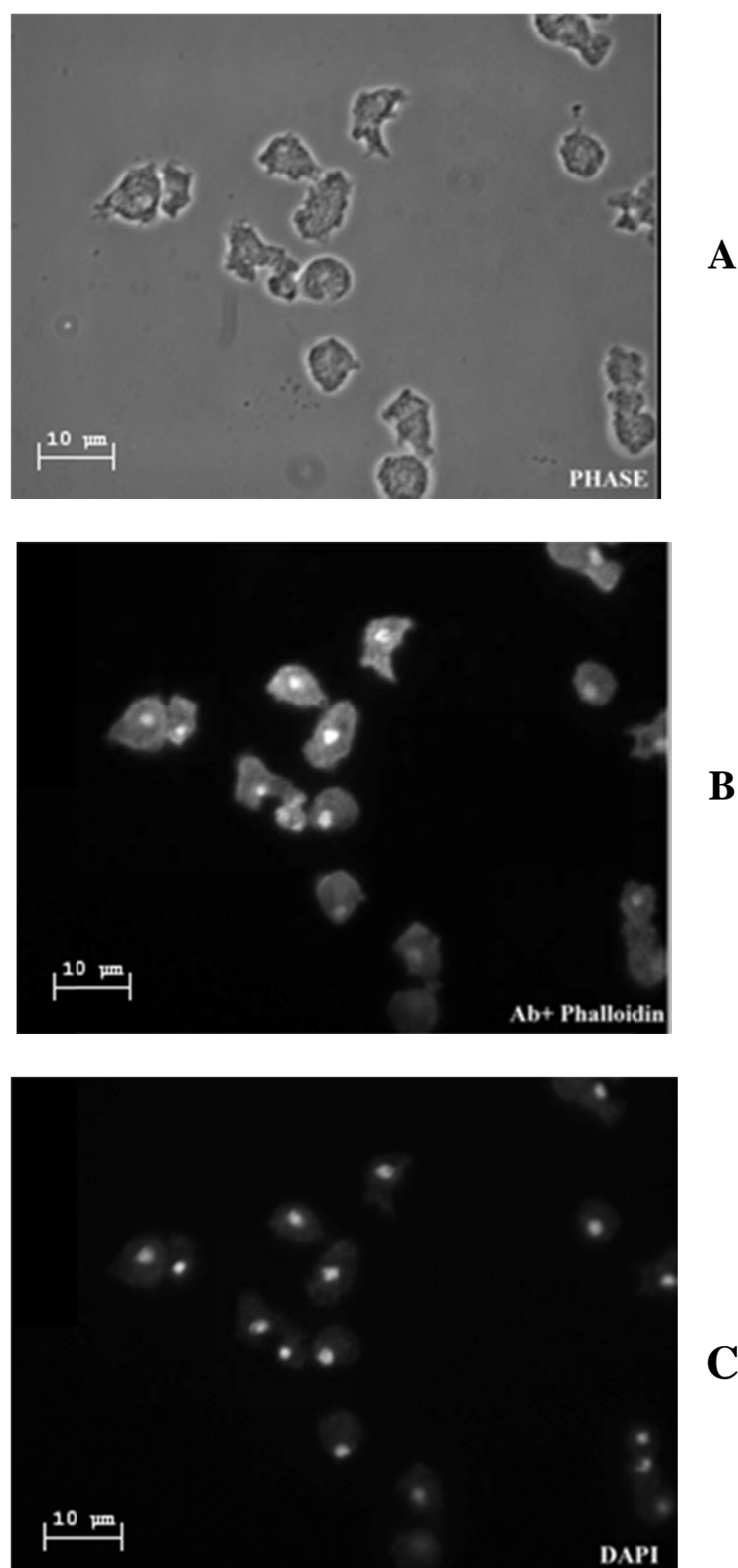


Figure 23: Immunofluorescence with Arp 5 antibody on fixed cells. (A) AX2 cells in phase contrast, (B) fluorescence after treatment with antibody and phalloidin, (C)

fluorescence after incubation with DAPI.

A knockout construct was made from the Arp5 genomic DNA as summarized in Figure 24. The construct contained the blasticidin resistance cassette between two homologous DNA pieces from the 5' and 3' ends of the coding region. The construct was linearized and transformed into AX2 wild type cells. The resistant cells were tested with PCR for replacement of the native gene. But no gene disruption could be detected. Because of the limited time left for the experimental part of the thesis, the knockout approach was not pursued further.

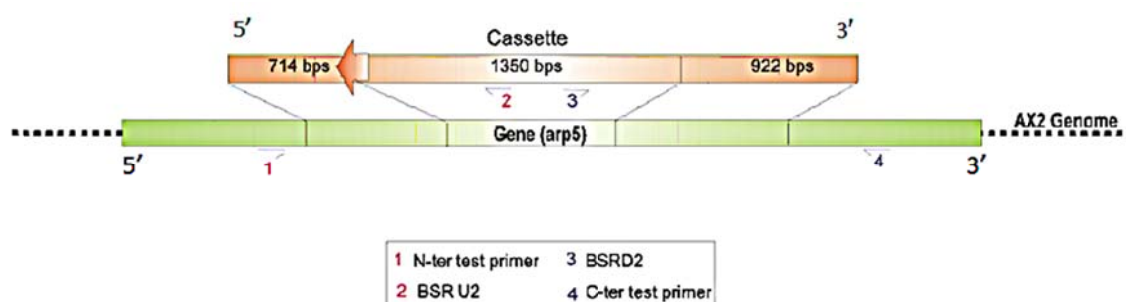


Figure 24: Schematical representation of the Arp5 knockout strategy. Cloning of the knockout vector followed standard gene replacement procedures. Essential is the permanent loss of an integral piece of the endogenous target gene. Therefore, the vector insert contained at the 5' and the 3' ends homologous regions with a length of 714 bp and 922 bp, respectively, flanking the blasticidin resistance cassette. In the case of a double crossover and the following insertion of the blasticidin cassette, the middle part of the gene would loop out and be replaced by the cassette. With this approach revertants are impossible. The detection of a successful gene replacement can easily be detected by PCR with selected primer combinations (in the scheme primers 1-4). Only after gene disruption the primer combinations 1+2 and 3+4 would lead to a PCR product.

3.11 Actin related protein 6 (Arp6)

The second nuclear Arp of interest for the current study was Arp6, which is associated with the SWR1 chromatin remodeling complex. The information that is known about it comes mainly from studies on yeast. As in the case of Arp5, a region unique to Arp6 was chosen. Polyclonal antibodies were generated against Arp6 amino acids #273-361 that had been purified as a Glutathione-S-transferase fusion protein.

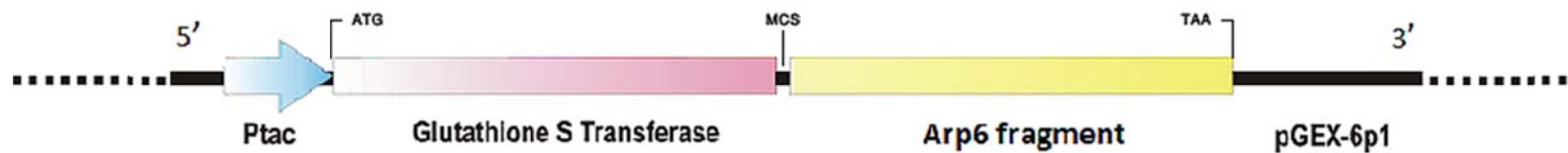


Figure 26: The scheme of the construct of Arp6 fragment fused to GST. The PGEX-6p1 vector forms the backbone, the *arp6* fragment was cloned via the BamH1 and SmaI sites. The fragment is cloned in frame with the Glutathione S transferase at the multiple cloning site (MCS) and TAA is added to the end of Arp6 fragment. The insert was sequenced and found to be without errors.

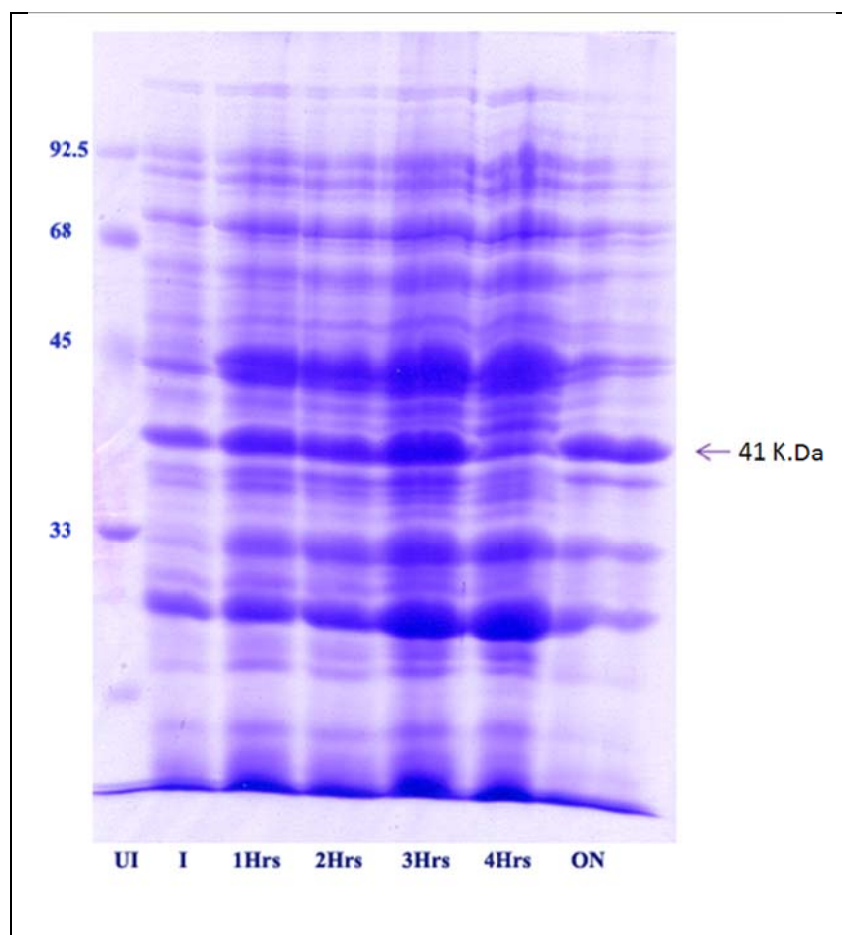


Figure 27: Expression pattern of the Arp6 fragment fused to GST. UI stands for uninduced while I stands for induced and ON for over night. 1Hrs, 2Hrs, 3Hrs, 4Hrs represent expression after 1, 2, 3, 4 hours after induction. The figure shows that expression after 3 hours post-induction gives maximum yield of the protein. The arrow points to the protein band with the expected size (~41 kDa). The nature of the protein was confirmed by mass spectrometry. Due to time limitations the recombinant protein was not characterized further.

A GFP construct was made by fusing the GFP DNA in frame to the 5' end of the Arp6 coding region. In the vector pDGFP-Xa-mcs-neo an Act15 promoter allows high expression levels. The construct was transformed into AX2 cells and clones that lit up green under the fluorescence microscope were used for further analysis. The positive clones were observed under different channels in the fluorescence microscope. The green labeling of the nuclei confirmed that Arp6 is a nuclear Arp.

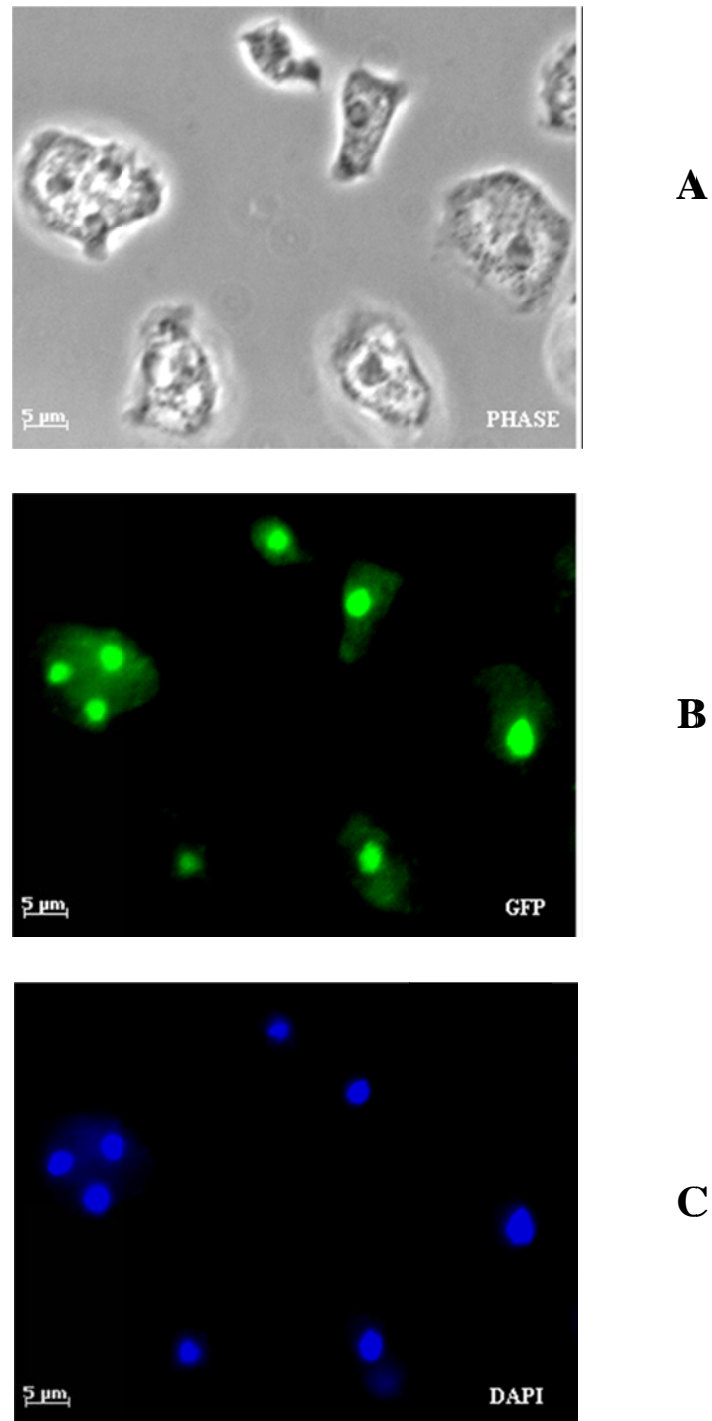


Figure 28: Localization of GFP-Arp6 in AX2 wildtype cells. (A) Phase contrast microscopy, (B) Arp6-GFP fusion protein is localized to the nucleus, (C) fluorescence after incubation with DAPI.

A knockout construct was designed essentially as described for the gene disruption of Arp5.



Figure 29: A pictorial description of the Arp6 knockout strategy. The homologous *arp6* 5' region contained 443 bases (#100 - 542), the 3' region 555 bases (#844 - #1398). In between them was the blasticidin resistance cassette. Through double cross over the blasticidin cassette would replace the original Arp6 gene. The positive clones were confirmed by PCR using diagnostic primers. The primers were designed in such a way that a PCR product from a primer 1+2 as well as a primer 3+4 combination proved the gene disruption.

4. DISCUSSION

4.1 Actin and protein evolution

D. discoideum, a model organism especially for studies on directed cell migration and leukocyte-like amoeboid movement, contains a large number of actins and actin-related proteins. Therefore, it was a prerequisite for further in vivo and in vitro studies to employ bioinformatic approaches. The large number of actins in *D. discoideum* was organized into three groups (Joseph *et al.*, 2008). The genomic locations of the genes have been identified and evolutionary trees shed light onto gene duplication events. An evolutionary tree involving all the members of the *D. discoideum* actinome was drawn to identify the relation between the members. An alignment of protein sequences between the identical actins and Arps was done to identify shared motifs and the structure of the identical actins was modeled based on the known structure of muscle actin. Structures were also modeled for the actin-related proteins Arp1, Arp2, Arp3, Arp5, Arp6, Arp11 and filactin. Domains and motifs of the members of the actinome were also studied from the aligned sequences, and it became very obvious that the modeled structures of putative actin-like proteins differed significantly. Sequence signals like the nuclear localization signal, nuclear export signal, protein export signal (SignalP), trans membrane signal and PEST signals were analyzed as well. Actin homologs across several taxonomic groups and the presence of Arps in several representative organisms showed on the one hand a drastic conservation during evolution. However, several other so-called homologs turned out as so different that one hesitates to use the usual designation 'actin-like' proteins. In parallel, first analyses on the actin-related proteins Arp5 and Arp6 were started to combine the knowledge gained from bioinformatic and experimental approaches.

4.1.1 Actinome

Actin is found in abundant amounts in nearly all organisms. The protein is highly conserved throughout evolution which suggests that at least the conventional actins were apparently under huge pressure in terms of structure and function. *D. discoideum* has a large and rather confusing number of actins, so the bigger picture was unclear until the actinome was studied in detail in the current project. As an outcome of this project, all the members of the *D. discoideum* actinome have been classified and organized. This includes the identical (conventional) actins and the non-identical actins of *D. discoideum*, the actin related proteins and the few pseudogenes. Comparative genomics revealed vital information on the occurrence

and multiplicity of actins and Arps in the organisms analyzed. It also gave away information on the conventional actin homologs of *D. discoideum* in other sequenced organisms, also projecting the key residues in their structural motifs. Fine structure analysis was done to the five key motifs of *D. discoideum* actinome. This exposed the structural, functional and evolutionary conservation and significances of these residues. Structural models were generated for key *D. discoideum* members of the actinome and were placed side by side for comparison.

4.1.2 Actin phylogeny

Mapping the *D. discoideum* actins on chromosomes 2 and 5 revealed the closeness and proximity at which these genes localized on the respective genomes. Phylogenetic maps were drawn with the DNA of the identical actins, which shows the duplication events that gave rise to this multiplicity of actins in *D. discoideum*. A tree with all the members of the *D. discoideum* actinome in it shows the phylogenetic relation between these genes. Especially interesting is the location of act8, which is at the tip of a series of duplication. This confirms the hypothesis that act8 is the most successful actin in *D. discoideum* and is the end product of multiple rounds of duplication. Act8 has been aligned against the Arps from *D. discoideum* to highlight the insertions and deletions. *D. discoideum* Act8 homologs from all available organisms were used to make a phylogenetic map. It is obviously not a perfect tree, as it suffers from non-availability of taxon-sampling but it shows an approximate relation between the different organisms.

4.1.3 Actin related proteins

Arps are believed to have originated in some organisms well before the evolution of *D. discoideum* as seen from the Fig. 2. Which explains why the Arps (represented in red) do not branch off from any *D. discoideum* actin and they do not give rise to any *D. discoideum* actin. They must have taken their separate track way before the evolution of *D. discoideum* and should have performed distinct functions as supported by the Arp anno paper findings. This must also be true with the two orphan Arps: Act31 and Act33. As depicted in figure 2, the orphan Arp Act32 has originated from and among the *D. discoideum* actins.

The Arp proteins were first classified in *S. cerevisiae* as Arp1 to Arp10 are based upon decreasing levels of sequence similarity to actin. Arp11 was later identified in the human

genome. Apart from this, there exists a family of orphan Arps. The *S. cerevisiae* genome encodes all but Arp11, whereas Arp7, Arp9 and Arp10 are specific to yeast. *T. rubripes* having the most compact genome, has just two Arp genes (Arp3 and Arp5). Hence these two Arps (Arp3, Arp5) can be also considered as the most essential Arps among the list of all the Arps. Another surprising thing which is obvious when one looks at table 8 is that Arps: Arp1, Arp2, Arp8, Arp11 have been known to have clinically relevant mutations but both Arp3, Arp5 don't. But this reflects just our current knowledge.

4.1.4 Structural evolution of actin

The actin molecule is more or less bi-symmetrical, with two sub-domains being symmetrical to the other two sub-domains. It is believed that before the evolution of actin there was an ancestor molecule which had only two domains but this molecule more likely underwent duplication to give rise to a protein with four domains. This ancient protein must have given rise to the precursors of actin like hexokinase, the Hsp70 family, other sugar kinases and prokaryotic cell cycle proteins such as MreB, FtsA and StbA. These molecules which are the structural homologs to actin have given rise to the present day actin structure.

4.1.5 Convergent evolution of actin

The actin structure is almost identical along a wide spectrum of organisms ranging from lower to higher organisms. The reason should be that there has been little room for change in a molecule which forms the building blocks of the cell. Even prokaryotic homologs to eukaryotic cytoskeletal proteins like MreB, CreS, Cren-1, FtsZ resemble the actin structure from eukaryotes. MreB even shows weak sequence identity with actin and has strong tertiary structural resemblance. It looks like actin got its structure by literally converging at the structure which it has today, a nearly perfect design and is not susceptible to any further structural change. As one could think of, buildings can come in all shapes and sizes but bricks are fairly standard.

4.1.6 Actin's structural motifs

As seen in Fig. 6, among the residues of the five motifs the size of the individual letters imply their significance at that position. Glycine is conserved in many places. Being the smallest

amino acid it occupies sites that have a high space constraint. Also being a neutral amino acid it occupies spaces where a charged amino acid would be troublesome. The next most conserved residue is proline, which is rather hydrophobic and without a charge. Thus, the neutral proline would help to get over charge constraints, and due to its hydrophobic nature it occupies sites that are more inclined towards the hydrophobic core of the protein. Proline has a significant role to play in the structure of any protein, as it provides a kink to the helix backbone. Aspartic acid is the next conserved, as there is certainly a need for a negatively charged amino acid at that position. Glutamic acid is the only other negatively charged amino acid but has a larger side chain than aspartic acid. So having aspartic acid at that position might be also due to size limitations.

4.1.7 Signals and motifs

The *D. discoideum* actinome was analyzed for the presence of various motifs and signals. SignalP looked for protein export signals and absence of which confirms that none of the members of the actinome are secreted out of the cell. TMHMM scanned the actinome for presence of transmembrane helices; the absence obviously shows that none of them are transmembrane proteins. PEST signals were also looked at, which is a signal for protein degradation. Arp1, Arp5, Arp8 and Filactin were found to harbor PEST sequences. It is known from previous experiments in our lab that Filactin is a highly unstable protein. Nuclear localization signals were mapped to be present in Arp5 and Arp8, which correlates with experimental observations in other organisms. Nuclear export signals were identified in almost all the organisms including those with identifiable NLS. As most proteins tend to move into the nucleus, NES plays a vital role in throwing them back to the cytoplasm. In the case of Arp5 and Arp8 the availability of both NLS and NES shows the dynamics with which these proteins are exported and imported to the nucleus (Feng *et al.*, 1999; Liku *et al.*, 2005). Apart from that, Arp5 and Arp8 have PEST sequences and are known to be less stable. The stability of these proteins could also be a deciding factor which controls the nuclear functions of these proteins.

4.1.8 Arp5 and Arp6

A region of Arp5 protein was expressed in *E. coli* successfully but it showed a thin upper band up on the Coomassie blue and the Western blot. Upon affinity purification the antibody

became more specific recognizing the Arp5 protein in an AX2 lysate. Arp5 was found to be localized in the nucleus as observed by immunofluorescence microscopy. This agrees with the observation in yeast; where Arp5 is part of the chromatin remodeling complex Ino80(Chen *et al.*, 2011; Udugama *et al.*, 2011). Furthermore, this is in concurrence to the bioinformatics procedure of nuclear localization signal prediction (NLS) that identified the sequence (NLS: KKKQR QLKSM KDGRL AQKRKR) which targeted Arp5 to the nucleus. Arp5 also has a nuclear export signal (NES: 323-ILALKTT-329) which exported Arp5 out of the nucleus. The presence of both NLS and NES is essential to regulate the dynamics of Arp5 protein present at the nucleus at a given time. A knockout experiment was also done with Arp5. The non-availability of knock out transformants could be either due to an innate problem that the gene is very essential for *D. discoideum* and is lethal when knocked out. Or it could be due to technical limitations.

In contrast to the ablation of Arp5, the disruption of the Arp6 gene in *D. discoideum* was possible. This could be explained by looking at the *T. rubripes* genome. *T. rubripes* is known to have the most compact genome, i.e. it possibly has only the most essential genes. Arp6 is not present in *T. rubripes* which suggests that it might be not as important as Arp5 which is one among the two Arps that *T. rubripes* harbor.

Transformants which expressed a GFP-Arp6 fusion protein showed green fluorescence in the nucleus. This confirms observations in yeast that Arp6 is part of the chromatin-remodeling complex SWR1(Deal and Henikoff, 2010; Bao and Shen, 2011). Though Arp6 does not have a predictable NLS sequence, its location in the nucleus needs another explanation. Arp6 could have a protein structure signal on the surface of the protein which would target Arp6 into the nucleus, in contrast to the presence of a NLS in Arp5.

4.1.9 Disease mutants of actin

ActC mutations have not been found in *D. discoideum*, which indicates that these are critical mutations and would have lethal effects on the organism. ActA mutations are relatively common in *D. discoideum* actinome, especially R183G and R183C mutations. R183W (Miller and Trybus, 2008) is a well-known mutation which causes severe ailments in humans. However this mutation is not found in *D. discoideum*; which highlights the negative impact tryptophan has at the position, against glycine and cysteine. Though there are other ActA mutations that are available in *D. discoideum* actinome these are not deleterious. These

replacements may be compensated due to other changes brought about to the structure of the molecule as a whole. With tools for expressing mutant actins in *D. discoideum* available now (Noguchi *et al.*, 2007), more active research in this field can be expected in the coming years.

4.1.10 Post-translational modifications in *D. discoideum* actins

There has been a selective advantage which has prevented the elimination of the redundant actins. Provided with the fact that *D. discoideum* actins differ at their upstream regulatory regions, this type of selection could be due to co- or posttranslational modifications which happen only at a specific time of development or at a distinct sub-cellular location. A number of post-translational modifications of actin in *D. discoideum* have been reported. There are acetylated and non-acetylated actins in a *D. discoideum* homogenate (Rubenstein and Deuchler, 1979; Rubenstein *et al.*, 1981), actin could be acylated in vivo with palmitic acid (Stadler *et al.*, 1985), but under certain environmental conditions actin is reversibly tyrosine-phosphorylated (Schweiger *et al.*, 1992; Howard *et al.*, 1993; Jungbluth *et al.*, 1994; Jungbluth *et al.*, 1995). Specifically during spore formation in the later stages of development tyrosine-phosphorylated actin forms bundles and tubes which get disintegrated on the onset of renewed germination (Sameshima *et al.*, 1994; Kishi *et al.*, 1998; Kishi *et al.*, 2000; Sameshima *et al.*, 2000; Sameshima *et al.*, 2001; Sameshima *et al.*, 2002). A recent report by Korn and coworkers shows that *D. discoideum* actin that was phosphorylated at Tyr-53 had an increased critical concentration, which is a greatly reduced rate of polymerization and a negligible nucleation activity (Liu *et al.*, 2006). Therefore, timely presence and correct localization of the appropriate tyrosine kinase might trigger this putatively co-translational modification. In such cases a tightly regulated expression would guarantee a highly efficient modification which is in fact responsible for a strong selective pressure.

4.1.11 Unique characters of actin

Actin is a slowly evolving but a highly expressed and a highly conserved protein. Other characteristics of a protein like actin can be seen from various references.

- A slowly evolving protein like actin accumulates mutations at specific locations in the protein; whereas a fast evolving protein will accumulate mutations all over the protein (Toth-Petroczy and Tawfik, 2011).

- A highly conserved protein like actin does not welcome mutations as it has many binding partners, in other words it has multiple binding/regulatory roles(Toth-Petroczy and Tawfik, 2011). So it is forced to preserve its structure and maintain its binding sites.
- The duplicates of a fast evolving protein tend to attain new functions (neofunctionalization), so that the derived genes can no more be recognized as duplicates. But the duplicates of a slowly evolving gene like actin are mostly ancient(Zhang *et al.*, 2003).
- When slowly evolving proteins like actin get duplicated by gene and genome duplication events they will be preserved and will continue their supplementary regulatory roles(Blanc and Wolfe, 2004; Seoighe and Gehring, 2004). So the amount of regulatory proteins increment with duplication events, as compared to other protein types. Consequently an organism gets more and more complex due to macro-mutations like duplications. This is in-line with the observed phenomenon that organisms have increased in complexity during the course of evolution.
- In terms of amino acid composition (at least in the folds); the slowly evolving and highly expressed proteins like actin tend to mimic the thermophiles(Cherry, 2010), by possessing an higher number of stable folds.
- As seen with actin there is a negative correlation between evolutionary rate and expression level(Drummond *et al.*, 2005). The slower a protein evolves the lesser it is susceptible to harmful mutations that cause misfolding. This misfolding and possible toxicity are strong selection criteria, which result in slower evolutionary rates but higher stability (Drummond *et al.*, 2005).
- Similar to actin a protein which is expressed in multiple tissues evolves slower than a protein which is expressed in a single tissue or only a few tissues(Kuma *et al.*, 1995; Hastings, 1996; Duret and Mouchiroud, 2000), because a tiny mutation is enough to make it incompatible in another tissue type. But if it is expressed in fewer tissues, then chances of it becoming incompatible is greatly reduced.
- Housekeeping proteins like actins obviously have to be conserved as they are less susceptible to change(Jordan *et al.*, 2002), because any defect in them could ruin the whole cell.
- **Moon lighting in actin:**if a protein performs multiple functions by binding to several other proteins then its behavior is known as moon lighting(Jeffery, 1999, 2003, 2004;

Gancedo and Flores, 2008; Huberts and van der Klei, 2010). Actin is known to exhibit moon lighting, as seen from the long list of actin binding proteins, which are about 285 in number.

4.2 *D. discoideum* and cellular evolution

D. discoideum is a good model to perform research at the organismic, cellular, and molecular levels. It is used to study cell differentiation, chemotaxis and programmed cell death, it is also used to study other aspects of cell development like cell sorting, phagocytosis, pattern formation, motility, signal transduction etc. It is believed that *D. discoideum* evolved after the plant-animal split but before the fungi-animal split (Eichinger *et al.*, 2005). Some unique characteristics of *D. discoideum* are discussed in the sections below.

4.2.1 Motility

D. discoideum exhibits both chemotaxis and phototaxis by moving towards e.g. cAMP and light, respectively. But *D. discoideum* is not the first organism to have evolved these functions (Faguy and Jarrell, 1999) as organisms lower to *D. discoideum* are also known to exhibit these behaviors. For example: *E. coli* moves towards glucose and away from benzoate and *Rhodobacter sphaeroides*, *Halobacterium salinarum* etc move towards light.

Motility is also found in organisms lower than *D. discoideum* and slime molds. But these lower organisms do not exhibit the 'amoeboid movement', they move by using structures like cilia, flagellae etc.. So actin based amoeboid movement was supposedly first evolved in a precursor of *D. discoideum*, which then was adapted by a variety of higher organisms including human. The human leukocytes exhibit behaviors like amoeba which include motility, phagocytosis etc.. But unlike *D. discoideum* there are only six major actin isoforms in the human genome. Analysis of the actinome of *D. discoideum* provides vital clues supporting actin evolution in *D. discoideum*'s precursor. Fig. 9 shows an evolutionary tree drawn using, non-identical proteins of the *D. discoideum* actinome. The tree suggests that through repeated rounds of divergence, Act8 has emerged as the most evolved member of the *D. discoideum* actinome. Surprisingly it also happens to have 17 identical protein copies. This proposes that Act8 is the most efficient actin in the *D. discoideum* actinome in terms of structure and function. This should also be the reason why it has been replicated and retained by *D. discoideum* multiple times. A sequence homology search of the human actins against the actinome of *D. discoideum*, illustrates Act8 as the best hit. This only proves that the highly evolved human actins are homologous to the identical actin members of *D. discoideum* and not to any other member of the non-identical actins. This emphasizes the pivotal role played by an

ancestor of *D. discoideum*, which has evolved the actin genes through repeated rounds of duplication. Incredibly, the process of evolution of actin has been frozen in time in an organism like *D. discoideum* which argues against any serious pressure from the environment, consequently it had not to take the austerity measures to get rid of the extra actin genes. Apart from this, the absence of actins in organisms lower than *D. discoideum* like the bacteria also adds support to the belief that primary actin evolution happened in a precursor of *D. discoideum*.

When *D. discoideum* made actins copies by duplication, genetic load would not have been a problem because, the primary gene performed the original function and the actin copy underwent minor modifications to become a new actin gene. By this way the multitude of actins were formed by repeated rounds of duplication and modifications. When an actin was no more useful to the cell, it began to dissipate by accumulating random mutations or became a pseudogene. In course of time this pseudogene could also be deleted from the genome. Presently there are six pseudogenes in *D. discoideum* actinome.

4.2.2 Multicellularity

Multicellular organisms have evolved independently at least 25 times from unicellular precursors (Bonner, 2000; Kaiser, 2001), if one defines biofilms as 'multicellular'. Fossil evidences show that precursors of cyanobacteria were the first multicellular organisms. Multicellularity through cell adhesion can be found in organisms as low as the bacteria *Staphylococcus*, but these multicellular organisms do not undergo cell-differentiation. *D. discoideum* on the other hand does exhibit cell-differentiation.

4.2.3 Cell differentiation

D. discoideum aggregates undergo differentiation producing both reproductive spore cells (psp) (80%) and vegetative stalk cells (pst) (20%) (MacWilliams HK, 1979). With specific markers one can detect even more tissue-like regions (among them upper and lower cup cells or disc cells in an emerging fruiting body). This cell differentiation procedure may be treated as a first step in the evolution of tissues which play specific roles in higher organisms. From this point of view, *D. discoideum* can be considered as being one among the first group of organisms to exhibit cell differentiation.

Once differentiation occurs among the cells, a certain set of genes are expressed in one cell type, while another set of genes are expressed in a different cell type. By taking a closer look at individual genes one would find that some isoforms of a gene are expressed in certain

tissues while other isoforms get expressed in other tissues. In terms of actins, human actin Act-c is expressed in the heart, Act-a is expressed in striated muscles and Act-b gives rise to the cytoskeleton of all cells and is also found inside the leukocyte cytoplasm.

4.2.4 Cheaters

Even among the differentiating *D. discoideum* cells, there could be individual cells which act with selfish interest by propagating themselves. These *D. discoideum* cells are called as 'cheaters'(Crespi, 2001; Fiegna and Velicer, 2003; Ackermann and Chao, 2004; Travisano and Velicer, 2004). Cheaters are in some ways similar to the cancerous cells of higher animals, which divide incessantly without considering the neighboring cells (lack of contact inhibition). Human cells have tumor suppressor proteins to hold back these fugitives, but it will be interesting to know if *D. discoideum* has tumor suppressors or other similar mechanisms to control the cheaters.

4.2.5 Tumor suppressors

A search in the dictybase using the term 'tumor suppressor' returns eleven genes which are homologs of human tumor suppressors like PTEN, TUSC4, LATS etc.. This confirms the assumption that *D. discoideum* had developed mechanisms to suppress the cheater cells. Being one among the first group of organisms to develop cell-differentiation, *D. discoideum* probably was also enforced to be a pioneer in evolving the prototypes of tumor suppressors. *D. discoideum* is as a good model organism in understanding the basics of this suppression system (Artemenko *et al.*, 2012), which could in turn help study the cancer suppression mechanisms of higher organisms(Iijima and Devreotes, 2002; Sherr, 2004).

4.2.6 Cell-cell adhesion

D. discoideum secretes cyclic AMP to attract neighboring cells to a central location. As the cells move towards the signal, they bump into each other and stick together by using developmentally regulated adhesion molecules which are in most cases glycoproteins. Gp150 is, for example, a membrane glycoprotein found in *D. discoideum* which is coded by the gene lagC(Dynes *et al.*, 1994; Wang *et al.*, 2000; Bolourani *et al.*, 2006). Null mutants of lagC are known to exhibit severe developmental defects.

Loss of cell-cell adhesion molecules is also the reason why cancer cells spread to other parts of the body. This process of metastasis seen during cancer progression could be curbed by studying membrane glycoproteins like Gp150 in *D. discoideum*.

4.2.7 Phagocytosis

D. discoideum is one among the lowest organisms to exhibit phagocytosis. *D. discoideum*'s phagocytosis of bacteria is a good example to show how DNA containing organelles like mitochondria came to be part of the eukaryotic cell. These organelles would have once been free living bacteria. But during a intracellular phase of symbiosis they might have lost the ubiquitous housekeeping genes as a result of a large scale deletion event in their genome. This would have forced them to choose between parasitic / symbiotic lifestyle by depending on other organisms for survival. *Mycobacterium spp.* are a good example of bacteria which lost part of their genome and opted for a parasitic lifestyle. Apart from bacteria being phagocytized by amoeba, recent reports of amoeba farming bacteria (Brock *et al.*, 2011; Palmer, 2011) only adds more value to this belief.

Genomic studies in *D. discoideum* have shown that it has maintained many of its ancestral genome diversity as compared to plants and animals (Eichinger *et al.*, 2005). But proteome based phylogeny demonstrated that after the plant-animal split, the amoebozoa have diverged away from the animal-fungal lineage (Eichinger and Noegel, 2003). From this point of view *D. discoideum* is a living relic together with the coelacanth and lungfish. Studying the evolution of *D. discoideum* actinome with its large repertoire of actins and Arps motivates and helps one to better understand the big picture of evolution and thereby attempting to explain how all the organisms have come to exist through evolution.

4.3 A sketch on the evolution of organisms

4.3.1 History of evolution

4.3.1.1 Neo-darwinism

Also called as 'modern synthesis', neo-darwinism is a combination of both classical darwinism and the mendelian genetic theory. Modern synthesis has been widely accepted, but its views on speciation are still under debate. Neo-darwinists insisted that speciation happens by gradual accumulation of small genetic changes over time. According to them this also includes gradual changes to the morphology over time, as one species gets converted into a narrowly distinct species. With this being a theory, confirmation for these events which happened during evolution over extended periods of time could be done only by studying the fossil remains. And the fossils tell a different story.

4.3.1.2 Punctuations in fossils

In spite of the neo-darwinian claim of gradualism, fossil evidences have not shown gradual stages of evolution from one species into its closely related species. Evolutionary trees drawn using these fossils show, rapid 'bursts' of speciation followed by long periods of 'stasis' where the species stayed unchanged. This model of evolution is referred to as 'punctuated equilibrium' and has been widely accepted (Gould, 1977; Mattila and Bokma, 2008). The absence of fossils which are missing links connecting two given species (interspecies missing-links), persuaded the non-darwinists to question the neo-darwinian views on speciation through fine gradations. And this successfully influenced the supporters of neo-darwinism; to accept and accommodate this fact by making changes to their theory.

4.3.1.3 Neo-darwinian speciation

With speciation being an occasional event, the neo-darwinians subsequently claimed that the genotypic and specially the phenotypic gradualism happen in isolation at secluded locations like an island separated by water barrier or far side of a mountain range which divided terrain acting as a land barrier (Hoskin *et al.*, 2005; Butlin *et al.*, 2008; Marie Curie SPECIATION Network, 2012). Neo-darwinists also claim that under these circumstances the fossils of the 'interspecies missing-links' could have been preserved, in remote sites away from where the common fossils are found, but would not have been unearthed until this day. With this being a theory, the presence of remote fossils has not been confirmed.

4.3.1.4 Morphological discontinuity

Darwin was aware of this problem and he believed that the 'morphological discontinuity' which was seen in fossils of his time, was because of incomplete fossil records. He hoped that as more fossils get unearthed in the future it would include interspecies fossils, which he believed would only add strength to his theory. But paleontology has come a long way ever since Darwin's days. Most of the fossil types unearthed now, are previously known ones and new fossil types are rarely being uncovered nowadays. Also the mathematical models in paleontology suggest that fossil records are nearly complete (Uyeda *et al.*, 2011).

4.3.1.5 Artifacts theory

Darwin and his supporters also believed in the 'artifacts theory'. According to which the fossil record was punctuated as it was full of artifacts; and all species were not equally represented, since many species did not get fossilized. They specifically claimed that organisms with soft bodies were not represented as they were difficult to be fossilized and to be part of the record.

The artifacts theory has been disproved recently with the excavation of soft bodied animal fossils in south-China (D.G. Shu and Chen, 1999).

4.3.1.6 Punctuations in taxa

Apart from the fact that fossils fail to support gradual speciation, the branching patterns observed in higher order taxa of both plants and animals also fail to support gradualism. Systematics and cladistics have shown that major features of 'body plans' (phyla) and their associated components arose not in the gradual way (Vergara-Silva, 2003). So biologists are on a constant search for alternatives to Darwin's gradualism all through the past one and a half century.

4.3.1.7 Alternatives to neo-darwinism

Though many alternatives to darwinism have been proposed, the famous ones are those which support non-gradual speciation. These include saltationism, mutationism, systemic mutation, quantum evolution, quantum speciation, hopeful monster hypothesis etc. They have gained some recognition after the emergence of the field of developmental biology (evo-devo) in the 1980s, but they lack the popularity which darwinism still holds.

4.3.2 Concepts in evolution

4.3.2.1 Micro-mutation and micro-evolution

Micro-mutations are small changes that happen to the genome of an organism like point mutations. This mostly brings about a change to the genotype and goes unnoticed (silent mutation), but rarely it does produce a phenotypic change. If this change results in a slight phenotype then it is called micro-evolution, but on the other hand if it is a quite remarkably altered phenotype then it is termed as macro-evolution. For example: a point mutation in hot spots like a developmental gene (Doebley *et al.*, 1997; Wang *et al.*, 2005) or at an active site of a protein, or a promoter of a gene has a better chance of resulting in a new phenotype, than one that occurs at an ubiquitous gene, the loop region of a protein or intron of a gene. So the outcome being a micro or macro-evolution depends upon the degree of phenotypic effect the mutation produces. Micro-evolution is generally referred to as a neo-darwinian mode of gradual evolution.

4.3.2.2 Macro-mutation and macro-evolution

Macro-mutations are large changes that occur in the genome of an organism like gross mutations, duplications, rearrangements, recombination etc, which results in a big change to the genome. These changes are generally reflected as a new phenotype but may also go

unnoticed by just contributing to the genotype. Here again, the scale of the phenotypic effect will determine if it is a macro- or a micro-evolution. Gene and genome duplications mostly result in a large new phenotype or macro-evolution. Macro-evolution being a relatively rare event, the evidences for its occurrence can be seen in the genomes of organisms which live today. For example: gene shuffling in antifreeze gene (Logsdon and Doolittle, 1997), genome duplications in fishes and other organisms (Pennisi, 2001; Cui *et al.*, 2006), macro-evolution of insect body plans (Ronshaugen *et al.*, 2002) etc.. Macro-evolution occurs abruptly (at random) and much less frequent than micro-evolution. This is the reason why one is not able to observe it, when one looks for it. But one can catch an event of macro-evolution if one would observe an organism continuously for many thousands of generations. So within a human life span one can observe macro-evolution only in organisms with very short generation times, like the microbe *E coli* (Lenski, 2004; Lenski, 2006) with a generation time of only 20 minutes or the viruses (Iyer *et al.*, 2006; Holmes, 2007; Simonsen *et al.*, 2007). Macro-evolution can also occur due to unanticipated (sudden) expression of silent micro-mutations (genotypic changes) that have been silently building up over successive generations, known as hidden evolution. Hsp90 is a good example for this (Fostinis *et al.*, 1992; Rutherford and Lindquist, 1998; Wagner *et al.*, 1999; Queitsch *et al.*, 2002). Under stress it releases the expression (phenotypic expression) of a variety of mutations in the developmental control proteins which had been accumulating until then.

4.3.2.3 Fast and slow evolution

Successive and rapid micro-evolutionary events, happening in an organism or in its proteins is called fast evolution. If these events occur relatively rare and sluggish, then it is termed as slow evolution. Slow evolution can be witnessed mainly in organisms which are under positive or neutral environmental pressure. As a slow evolving organism which experiences negative selection pressure would have entered the endangered list or would have gone extinct.

4.3.2.4 Environment's role

The environment which is an external factor also plays the prominent role in shaping the evolution of organisms. Some organisms like the puffer fish have had such severe 'negative environmental pressure' (purifying selection) that they lost the extra genes by deletion and have resulted in having compact genomes. On the other hand fishes in the arctic and antarctic have evolved antifreeze proteins which give them an advantage over the environment or in

other words they are under 'positive environmental pressure' (adaptive selection). There exists a third type of external pressure which is the 'neutral environmental pressure' seen in organisms like the lungfish and coelacanth. This neutral pressure allows them to continue having their rudimentary hind legs instead of fins. It is also due to this neutral environmental pressure that *D. discoideum* has a myriad of actin genes.

4.3.2.5 Species branding factors & 'hypothesis of species branding'

Species branding factors (SBFs) are the molecular factors which spear head the process of speciation, these factors tag a given species there by preventing it from reproducing with any other species. Some of the proteins involved during fertilization are good examples for species branding factors. These proteins are part of all the three steps of fertilization; thus they ensure species-specificity at multiple levels:

1) Chemotaxis of sperm occurs at a species specific manner and has been documented in several species of echinoderms, molluscs, urochordates, and cnidarians (Miller, 1985; Yoshida *et al.*, 1993). The protein 'resact' attracts the sperm migration in sea urchin, *Arbacia punctulata* (Ward *et al.*, 1985).

2) Sperm and egg interaction is also species specific. Surface protein of the sperm such as 'bindins' attach with receptors such as 'EBR1' on the egg's vitelline membrane seen in sea urchins (Kamei and Glabe, 2003). In the mouse sperm a trans-membrane protein called 'fertilin', enables its binding to the protein 'integrin' on the egg plasma membrane (Yuan *et al.*, 1997; Alberts, 2002; Evans and Florman, 2002).

3) Fusion of the nuclei from the egg and sperm is also species specific, which is brought by 'fusogenic' proteins similar to 'HA protein' of influenza virus (Plotch *et al.*, 1999; Hamilton *et al.*, 2012) and the 'F protein' of Sendai virus (Bagai *et al.*, 1993). Bindin and fertilin are also believed to promote cell fusion in sea urchins and mouse, respectively.

Nuclear gate keepers (NGK) (Capelson *et al.*, 2011) are proteins which are stationed at the nuclear pores where they prevent the entry of genetic material from another species through viruses. They are highly species specific and thus prevent breeding across species. Proteins such as bindin and NGK are probably the fastest evolving proteins and are believed to diverge rapidly when compared to other proteins of the organism. They brand a given species, prevent its interbreeding and enable it to diverge far from the parental species. They probably are also the molecular basis behind the classical definition of species; which says that a species does not interbreed with another species.

Prokaryotes which do not have a nucleus and so the NGK are also known to have evolved similar mechanisms to protect and brand the host genome. Archaeobacteria as well as eubacteria possess the restriction-modification system; (Wilson, 1991; Wilson and Murray, 1991) which chop any foreign DNA entering the host cell. They would also safeguard and brand their genome by methylating it at unique sites. Codon bias and GC content which varies between organisms play a vital role in branding, by providing the host DNA an upper hand over any foreign DNA that might have sneaked into the cell.

There are also other genetic factors which aid in species branding and speciation, like the 'interspersed repeats' (Jurka *et al.*, 2011). These repetitive DNA stretches get scattered within newly evolving 'copies of a gene'. They preserve the new gene from being replaced by the 'original gene' through homologous recombination. The original gene might exist as a paralogue in the same genome, or might enter from another individual through mating. The interspersed repeats thereby protect and enable the new copy to diverge by accumulating mutations and to evolve new functions

4.3.2.6 Quantitative traits

Variations in morphology are seen in members of any species that exists today (at least due to SNPs). These variations are necessarily not the products of micro-evolution, but are the quantitative traits brought about by polygenic effects. These traits are nothing but heritable 'fluctuations'. They never contribute to the gradual process of evolution because the population still remains as a single species, is interbreedable and the traits are mostly "reversible". They are the reason behind the variations seen in Mendel's peas; for example the height, which has been known to 'fluctuate between generations' and has not continued to increase through evolution. The different breeds of dogs may also be treated as quantitative traits, because of their ability to interbreed and the reversibility of certain traits. Though micro-evolution happens today, what appears like gradual evolution in most cases is actually heritable 'fluctuations' (quantitative traits) (Paran and Zamir, 2003; Mackay *et al.*, 2009; Plomin *et al.*, 2009), which are rather widespread. The effects of quantitative traits on paleontology, where classification is primarily done based on morphology, remains an open question. Especially some of the admired 'transitional fossils' (missing links) with fine-gradations in morphology might not be the products of micro-evolution, but be variants brought about by quantitative traits. Fossils of these variants which slightly differ in morphology, could enable the curators to arrange them as 'series of specimens' and to present them as evidence for the 'different stages of progression' of micro-evolution.

Obvious quantitative traits produce several morphological variations which tend to fluctuate. So it would be better not to use morphological variations as a criterion, when gauging the fine-gradations of 'micro-evolution' in the organisms that exist today. But it would be wise to use the "degree of incapacity to interbreed" (hybridize) caused by the divergence of species branding factors, as the chief criterion to measure the progress of micro-evolution.

4.3.2.7 Species concepts

1) Interbreeding evasion: From the classical definition, species are a group of organisms which are not capable of interbreeding with other species in the wild. During micro-evolution when speciation begins the newly formed variants mostly prefer mating with similar variants (inbreeding), than interbreeding with dissimilar variants. For example: two groups of *Drosophila pseudoobscura* were fed with two different food types for eight generations (Dodd, 1989). At the end of the experiment they preferred mating only with those individuals, who have consumed the same food type (similar variants), over mating with those that had consumed the other food type (dissimilar variants). In this sense *Drosophila pseudoobscura* consuming a certain type of fruit for successive generations, will avoid breeding with variants of the same species which have been consuming another type of food for several generations.

The interruption in breeding between morphologically identical successors of *Drosophila pseudoobscura* (reproductive isolation) may be due to factors such as the epigenic switches like methylation. They could isolate the 'new variant' of a species by avoiding it from mating with other members of the same species, there by initiating the origination of a new species. There is also the possibility that the reproductive isolation could be due to simple reasons such as preference due to body odor, caused by consuming different food types. In any case interbreeding evasion could be an initiating factor in the long run towards speciation.

2) Interbreeding in captivity: Some variants are capable of interbreeding in captivity, while they do not normally interbreed in the wild. These breeds have diverged such that they interbreed readily under lab conditions (captivity) but interbreed less often in the nature (wild). For example: interbreeding has been reported between several species of Darwin's finches which interbreed more often in the lab than in the wild. Because of their ability to interbreed these species can no more be considered as distinct species, but have to be regarded as different breeds (dissimilar variants) of a single species. After Darwin had left the Galapagos Islands, continued research on the Darwin's finches has observed that the different species of finches that were celebrated by Darwin do in fact interbreed with each other even

in the wild. *Geospiza fortis* has been reported to interbreed with *G. magnirostris* (Huber *et al.*, 2007), *G. fuliginosa* interbreed with *G. fortis* (Grant P. R., 1994), *G. scandens* interbreed with *G. fortis* (Grant and Grant, 2002).

Recent studies in molecular biology on Darwin's finches confirm that their beak morphology is the outcome of five different genes (polygenic) (Mallarino *et al.*, 2011) i.e. Bmp4, Calmodulin, TgfbIIr, β -catenin, Dkk3. This adds support to the view that the different species of Darwin's finches can be considered as mere variations and those that interbreed may actually belong to a single species.

Reports of the different species of Darwin's finches interbreeding in the wild, questions the very definition of species. The classical definition of species, as organisms which do not interbreed with other species in the wild, seems to be fragile and may need an update.

3) Interbreeding *invitro*: As speciation diverges the dissimilar variants further, they are capable of interbreeding only in the lab through *invitro* fertilization, and not even in captivity. This could be because of some 'species branding factors' which have diverged quickly. This could include the cell surface proteins which are needed for the cell entry process of the sperm. *Invitro* fertilization overrides these proteins by injection of the sperm directly into the cytoplasm. For example: sickle back fishes are found in the northern hemisphere where their different species vary in features like body size, body shape, presence or absence of the defensive armor, size & pattern of skeletal structures, etc.. More than 40 different species of stickleback fishes have been classified. Among them some species can interbreed through laboratory mating, while there are some other species which can interbreed only through *in vitro* fertilization (Bell, 1994; Chouard, 2010). Because the variants yield to interbreeding through *invitro* manipulations means that they have not diverged far enough to be called distinct species. Therefore, they can be considered only as 'un-established species'.

4) Fully branded species: With further deviation of the diverged variants, their 'species branding process' gets completed. As a result they can never interbreed either in the wild or in captivity and even with the help of *invitro* techniques like *invitro* fertilization. At this stage they may be authenticated as separate species or as stable morphological forms. These stable species might have the potency to establish them in large numbers and to leave behind ample fossils.

A stable species is one which has distinguished its species branding factors far-away from its species of origin. As widely accepted a species radiates into different morphological variants; when a new niche is born or when a former niche gets vacant, following a mass extinction

event or due to artificial selection. From the text above it is evident that, all these progenitors of a species like a few finch varieties and some sickle back fish variants may not be classified as different species because of their interbreeding behaviors; but could be considered as different morphological variants or subspecies. It could take a while until stable species emerge out from among these variants. These stable species are capable of leaving behind the ample amount of fossils required to be discovered through excavations in the future.

4.3.2.8 Prospects of micro-mutation

1) Micro-mutations in species branding factors (SBFs):

Micro-mutations happening in the genes of species branding factors (SBF) might result in the origin of a new sub-species. These new subspecies over time could speciate/ radiate in to new species. This process of speciation may employ the different species concepts by passing through different stages of speciation like inter-breeding evasion, capacity of inter-breeding only in captivity, capacity of inter-breeding only by *in vitro* means, establishment of stable species. These radiating new species get adapted to the environment by occupying a micro-niche.

2) Micro-mutation happening in non-SBF genes

- Most micro-mutations result in new genotypes. These new genotypes are primarily due to single nucleotide polymorphisms (SNPs) that are not known to exhibit a phenotype.
- Relatively rarely new phenotypes are produced by micro-mutations. This is the result of SNPs that are capable of exhibiting a phenotype. These SNPs could be associated with quantitative traits (QTL) which produce the different morphological variants or breeds.

4.3.2.9 Prospects of macro-mutation

Macro-mutations almost always result in macro-evolution by bringing about major changes to the genome which most likely would include the SBFs.

Case 1: Genome level duplication

Following a genome level duplication event, changes could have taken place to the SBFs. But because copies of the original SBFs are present; this will permit continued mating with the parental population, resulting in hybrids. With the huge incrementations to the genome produced as a result of the duplication, it would probably take at least a few generations for the organism to settle down into a relatively stable form. The mutant mating with the parental population would only assist the smooth occurrence of this process. Over time hybrid SBFs would be generated which would enable inbreeding only among the new hybrids. The

intensity of phenotypical change brought about by this event will determine if a new species, or a genus or order or class is produced. With new features not found in any of its relatives; the organism would sneak into a new ecological niche if possible into a new ecosystem, thereby adopting a new environment. This way fishes could have adopted the arctic or antarctic environments.

Case 2: Macro-mutation not involving SBFs

This case of macro-mutation results in vast differences in between the parental species and the derived species, as compared to the variation in morphology found among the members of parental species or the derived species. As not much change occurred to the SBFs, despite of vast morphological differences (produced by the non-SBF genes), these two species are capable of mating to produce offspring. In many cases the offsprings are sterile. For example: Tiger and lion though belonging to different species can mate to produce offspring; similarly horse and donkey can mate producing the mule, which is mostly sterile.

Looking at the punctuations seen in the fossils, one can imagine that macro-mutations mostly lead the vertical evolutionary process, while micro-mutations contributed mainly towards horizontal (lateral) evolution.

Micro and macro-evolutions from a paleontological perspective:



Figure 30: A depiction of micro- and macro-evolution in terms of fossil records.

4.3.2.10 Evolution in action

While the neo-darwinian views insist on evolution only through gradualism the other darwinian concepts like selection, adaptation, survival etc. are widely acknowledged as visible proofs for evolution happening even today. To quote a few examples of evolution in action: environmental selection either natural or artificial (the peppered moth *Biston betularia*) (Hart *et al.*, 2010), the survivor is the fittest and it can propagate (the cane toad of Australia) (Urban *et al.*, 2007), adaptation to the environment (Italian wall lizard *Podarcis sicula* evolution in the island of Pod Mrčaru (Losos *et al.*, 2001)(possibly macro-evolution),

the yellow bellied three-toed lizard *Saiphos equalis* of Australia which either lays eggs or gives live birth as an adaptation to coastal regions or inland area(Linville *et al.*, 2010), adaptive variations (Darwin's finches)(Petren *et al.*, 2005; Tebbich *et al.*, 2010), evolution of drug resistance in bacteria(Albrich *et al.*, 2004; Bennett, 2008), and viruses(Chen *et al.*, 2004; Yim *et al.*, 2006)(mostly saltation), etc.

4.3.3 Cambrian explosion

Cambrian explosion of life forms is biology's equivalent to the big-bang;(Levinton, 1992; Kerr, 1993; Nash, 1995; Thomas *et al.*, 2000) as major groups of animal fossils (evolutionary novelties) can be obtained from just this layer on earth, compared to any other layer. A vast variety of new 'body plans' emerged in the cambrian period. Aquatic animals belonging to about 50 distinct phyla appeared abruptly, which includes representatives of almost all the 35 phyla which live today. Apart from these, fossils show that at least 15 other phyla must have lived during the cambrian, which later went extinct. Surprisingly only those phyla found in the cambrian seas have evolved further into the land animals we see today, hardly any new phyla have originated ever since the cambrian era. Missing links which connect the cambrian phyla with fossils that predates the cambrian are hardly found. Which suggests that these phyla abruptly came into existence only during the cambrian period. Apart from this, missing links connecting these distinct phyla to one another have also not been found. The variety of fossils found in the cambrian is an explosion of life, provided the relatively meager range of fossils found in the pre-cambrian periods. The distribution of cambrian beds across different continents, emphasizes the fact that this was a global event. Cambrian explosion started at about 543 mya ('million years ago') according to the 'international sub-commission on cambrian stratigraphy' (2002) and occurred within a short period of 5-10 million years(Bowring *et al.*, 1993; Monastersky, 1993). In relative terms this episode took place in a fraction of time when compared to the 3.8 billion years of life on earth (eoaarchean era)(Mulkidjanian *et al.*, 2012). But the factors which caused this event to occur and to end are mysteries, which remain largely unanswered even today.

The extrinsic or environmental factor which caused this event are believed to be an influx of oxygen (Saltzman *et al.*, 2011) in the cambrian atmosphere; as seen from the traces of iron-oxide (Klein, 2005) found in the cambrian rock layers. This oxygen is believed to have been released through the aerobic bacteria and photosynthetic phytoplankton, cyanobacteria &

algae which had evolved before this event (Jackson and Cheetham, 1990). Oxygenation of deep oceans (Fike *et al.*, 2006; Robert and Chaussidon, 2006) enabled the origin of large bodied animals (Butterfield, 2009). The evolution of lungs and gills supported the increased requirement of oxygen by these large body forms; as compared to oxygen assimilation through the skin which was in use until then. The cooling of the surface of the earth by then is also believed to have supported the cambrian explosion.

Fossil evidences prove that very slow evolution had happened in the periods before the cambrian era and ever since the conclusion of the era (Arthur, 1997). So the only way to explain the explosive cambrian evolution and the swift appearance of phenotypic novelties is that: it had happened through fast and abrupt transitions, in other words by mega-evolution. In this work the author has attempted to explain cambrian explosion and mega-evolution using his evo-nhancer hypothesis.

4.3.3.1 Mega evolution / super evolution / enhanced evolution

Successive macro-evolutional events which result in a diversity of organisms within a relatively short period of evolutionary time is called mega-evolution or super-evolution. This process is believed to have involved rare macro-mutational events like gene/genome level duplications, polysomy, polyploidy etc. together with intermittent micro-mutational events which as a whole provided a huge amount of genetic variation to the organisms. The cambrian era of evolution is a typical example for super-evolution.

The huge variations in morphology seen among the cambrian fossils was most likely due to mutations in the developmental genes like Hox, DII, MyoD, Pax-6, Sog, Bmp-4 etc. (Levinton, 2008) which gave rise to the major body plans (phyla) that we see even today. A good example is the antenapedia mutant of *Drosophila melanogaster* (Abbott and Kaufman, 1986; Randazzo *et al.*, 1993; Ronshaugen *et al.*, 2002) which demonstrates that even small mutations could bring about a huge morphological variation such as the antenna getting replaced by leg-like organs. So on empirical basis there need not be hundreds of millions of years to pass by for a new body form to develop (Doebley *et al.*, 1997; Wang *et al.*, 1999; Wang *et al.*, 2005). Mega-evolution has the ability to generate radically new morphological novelties within a few million years.

But it is relatively easier for one to interpret mega-evolution or its smaller cousin macro-evolution as cumulative effects of gradual changes to morphology over eons of micro-evolution, by not paying any attention to the fossil-evidences.

4.3.3.2 Evo-nhancer hypothesis

To explain cambrian explosion which happened through mega-evolution, and their fossil evidences exhibiting punctuated equilibrium (Gould, 1977; Gould and Eldredge, 1993) one could propose a putative molecule called evo-nhancer. Evo-nhancer is most likely a protein which played the role of a master switch that enhanced the activity of other proteins which caused molecular events like duplications, mutations, recombination, deletions, polyploidization etc. These are caused by enzymes like recombinase, transposase etc. According to the evo-nhancer hypothesis, evo-nhancer is believed to be the intrinsic or biological factor behind all the rapid changes which took place during cambrian evolution.

Given the relatively slow pace of evolution seen from the fossils that lived prior to the cambrian explosion, it is highly unlikely that the evo-nhancer molecule had existed by that time. The protein evo-nhancer most likely came in to existence at the start of the cambrian era, due to mutations in an ancestral protein of the evo-nhancer. This mutation might have been caused by reactive oxygen species (Hsie *et al.*, 1986) found in the oxygen rich atmosphere which had existed at the start of cambrian era.

According to the presumed process evo-nhancer leads to active changes inside the cell, which leads to massive changes in the genotype and phenotype of the organisms resulting in mega-evolution. This active evolutionary process possibly occurred in individuals/tiny groups of a population, independent of other individuals/tiny groups and at random with the amount of change in between them varying. These individuals/tiny groups could be called 'transition-species' which left the ecological niche they had occupied till then and moved to colonize new niches, thereby 'adopting' an environment. This process of active evolution and colonization had repeated time and again resulting in cambrian life forms exploring new niches and colonizing the earth. In the middle of these repeating cycles, the evo-nhancer protein could have been lost in some individuals/tiny groups of a population, more likely due to deletion or mutation of the gene that coded for the evo-nhancer protein. These individuals/tiny groups known as 'pro-species' might have continued to stay in the niche they had colonized 'reaching fixation' and, failing to evolve into organisms higher than them. These pro-species would have got well adapted to their niche through micro-evolution and could have evolved over the years into the species we see today.

The process of mega-evolution which happened during the cambrian era occurred like a chain reaction leading to an explosion of life forms on earth. This process probably halted; when the evo-nhancer molecules were all lost by deletion or dissipation, and probably also because of some environmental factors. Evo-nhancer as of today, is assumed be an 'extinct molecule'. There is also a rare possibility that dormant versions of the evo-nhancer exist as pseudogenes in some life forms. After the cambrian explosion, mega-evolution slowed down in pace and micro-evolution and macro-evolution took over its place. One can assume that even in the periods before the cambrian era, organisms could have evolved only through micro-evolution and macro-evolution.

Though a variety of new phyla originated during the relatively short period of the cambrian era few phyla did originate in the periods prior to and since the cambrian, but this happened rarely. This probably happened by intermittent or sudden expression of the additive micro-mutations (to the genotype) spanning innumerable generations, or due to sporadic events of macro-mutation which gave rise to hopeful monsters (Page *et al.*, 2010).

Some scientists believe that Hox genes were the 'prime biological factors'(Grenier *et al.*, 1997; de Rosa *et al.*, 1999) behind the cambrian explosion. But if that would have been the case then the cambrian era would have ended with the extinction of the prime factor Hox. But Hox exists even today; this only suggests that Hox was not the main biological factor but that it might have stayed under the control of another regulator, which is assumed to be or like the evo-nhancer protein. Also, if Hox would have been the prime factor then the continued presence of Hox would have led to further mega-evolutionary events in the eons that followed the cambrian, which is not known to have occurred.

As a fact, mutations do not always result in the victorious 'hopeful monsters'(Theissen, 2006, 2009). Indeed, most mutation events result in the unsuccessful 'hopeless monsters'. But because mega-evolution happened swiftly and was more intense, the amount of successful mutants produced by this process was significantly higher, as indicated by the variety of cambrian fossils. The hopeless monsters obviously got filtered out through natural selection. Thus both mutation and natural selection influence evolution, which is aptly known as 'dual causation'(Stoltzfus, 2006).

4.3.3.3 Support for mega-evolution

Mega-evolution being an event which happened way back in time cannot be explained with living examples, and can possibly be explained only through analogies.

In vivo studies in recent years have come up with new varieties of plants which have been made by exposing them to mutating radiations, mutating chemicals etc. This process is followed by the selection of the desired variety of plants (Wright *et al.*, 2005; Burgess *et al.*, 2007) (artificial selection). This analogy shows that: if nature were to develop a new species in a similar way and in a relatively short span of time then these mutations should have happened through a master mutator like evo-nhancer, followed by natural selection.

Activation-Induced Deaminase (AID) mediated 'antibody diversification' and 'somatic hypermutation' is a good analogy to explain how fast and how much variation could be brought about within a single generation. During antibody generation especially the 'affinity maturation of B cells' (Teng and Papavasiliou, 2007; Weiser *et al.*, 2011) AID-mediated somatic hypermutations occur at the variable regions of antigen-binding coding sequences, resulting in thousands of slightly different immunoglobulins, which are followed by 'clonal selection'. AID is presently believed to be the master regulator of the secondary antibody diversification process. AID can be considered as a scaled down version of the evo-nhancer principle and is a living proof of super-evolution.

Huge numbers of genetic variations similar to those generated by AID have been successfully reproduced *in vitro* through a process known as 'directed evolution' (Turner, 2003; Rubin-Pitel and Zhao, 2006; Worsdorfer *et al.*, 2011). Random errors are introduced into the DNA of a given protein through PCR using an error prone DNA polymerase. This is followed by expression and selection of those mutant proteins which exhibits increased efficiency. The loop continues by expecting incrimination of efficiency with each loop, until the desired mutant protein is obtained. The analogy shows that mutations don't have to be necessarily harmful but that beneficial mutations do happen in reality, provided the right selection environment exists.

These analogies not only add support to the 'hypothesis of enhanced evolution', they also confirm it. By philosophizing logically if rapid evolution could take place in these cases, then in the presence of a master mechanism like evo-nhancement evolution would have certainly happened at a fast pace. It also shows that a cambrian organism would have possessed 'evolvability' due to the presence of the internal factor evo-nhancer.

4.3.3.4 Advantages of the evo-nhancer hypothesis

Evo-nhancer and its role in mega-evolution can be compared to Occam's razor; according to which "the simplest explanation for an observed phenomenon is more likely to be accurate when compared to more complicated explanations" (Thorburn, 1915). Consequently, evo-

enhancement would explain the observed facts of cambrian fossil evidences, punctuated equilibrium and mega-evolution in 'a much simpler fashion' as compared to many other theories. It even explains the molecular basis behind the morphological discontinuity seen in the fossils and elucidated by the punctuated equilibrium theory (Gould, 1977; Gould and Eldredge, 1993). It also can explain the mechanisms by which the different phyla came into existence and how they diverged to finally colonize the earth.

4.3.4 Resurfacing of past theories

The recent availability of genome information is making the evolutionary histories of many organisms to be re-written (Brown, 1999; Dalton, 2010). The advent of the genomic era is also generating new evidences to support former theories of evolution like lamarckism. In a debate between gradual versus punctuated change one could make an educated guess that darwinism is right in the sense that gradualism does occur often at the genotypic level. Saltationism is also right in its own sense that 'morphological discontinuity/punctuation' occur typically at the phenotypic level, seen as noticeable changes to the morphology.

4.3.4.1 Neo-lamarckism and epigenetic inheritance

Surprisingly, the discarded theory of lamarckism has also earned recognition in recent years as neo-lamarckism or soft-inheritance. According to which characteristics acquired during an individual's lifetime can be transmitted to their offspring in non-mendelian ways. Epigenetics (Casadesus and Low, 2006; Chahwan *et al.*, 2010; Banister, 2011; Chahwan *et al.*, 2011) explain how this non-genetic inheritance happens and describes the molecular events behind

it. The environment is known to affect regulatory proteins like hormones which in turn methylate the DNA at unique sites, thereby switching on and off the expression of specific genes. Though this DNA gets inherited the mendelian way by being bonded to it the modifications, e.g. methylation, are possibly passed on as well. Therefore, the genes are inherited either switched on or off which in turn influences the phenotypic characters of the new generation. So in the end, characters acquired during one's lifetime would be passed on to his progeny by epigenetic inheritance. It should be emphasized however, that basic research on epigenetics currently move with high speed on a sometimes meandering track; i.e. all current hypotheses about epigenetic inheritance should be taken with a grain of salt.

4.4 Natural history of earth(Cohen, 2012)

1) When the earth was formed, it was a ball of hot gases. Geological records dating the earth's earliest rocks indicate that earth is about 4.6 billion years old (the hadean period).

2)The hot earth initially must have been surrounded by a thick cloud of gases and steam which formed its atmosphere. This thick cloud might have been several kilometers high and because of it sunlight would have hardly reached the surface of the earth.

3) As earth started cooling, water vapor in the cloud had slowly begun condensing in to liquid water. This water must have got collected at the fissures on the surface of the earth. The clearance of the clouds must have initiated the formation of the earth's atmosphere. Chemical evolution most likely took place in the pools of water accompanied by lightning and radiation, leading to the formation of the basic building blocks of life like the self replicating RNA, amino acids etc.. This probably led to the formation of a proto-cell from polymers (coacervation theory). It is also discussed that the proto-cell arrived from somewhere in space through asteroid impact on earth.

4) Biological evolution started in the eoarchean era of the archean period, with the primordial life developing in to archaeobacteria. These bacteria used sulphur and other chemicals to gain their energy. Being thermophiles they lived in the hot waters & thrived in the anerobic environment. The paleoarchean era followed with the emergence of the first known oxygen-producing bacteria as indicated by the oldest among definitive microfossils. These aerobic-bacteria slowly raised the concentration of oxygen in the atmosphere thereby pushing the archaea to secluded anaerobic locations on earth like deep ocean vents and hot springs, where they survive till this day.

Cyanobacteria, phytoplanktons etc. evolved in the mesoarchean era in water bodies, which started the process of photosynthesis with the available sunlight. Stromatolites of cyanobacteria discovered in Australia are among the oldest macro-fossils found till now, they date back to 3.485 billion years ago (the archaean period)(Grotzinger and Knoll, 1999; Allwood *et al.*, 2006). They also released the byproduct oxygen, which made the earth's atmosphere rich in oxygen (orosirian era) and might also have lead to the formation of the ozone layer. The

ozone layer acted (still acts) as a shield protecting life forms from harmful UV radiations there by promoting the evolution of advanced life forms. Statherian era gave rise to the first complex single celled organisms, in other words the protists with nuclei (eukaryotes).

5) Meanwhile geological evolution was also happening in parallel with columbia being the first primordial supercontinent which led to the rodinia supercontinent (tonian era) and then the supercontinent of pangaea (permian era) which then split to produce the various continents that exist today. These changes happened as a result of the continuous motion of tectonic plates. This movement of earth's crust created the environmental niches much needed for the evolution of life like water pockets, lakes, mountains, valleys, seas, islands and oceans.

6) As earth was cooling further nearly all the remaining clouds condensed in to liquid water, leaving behind an almost clear atmosphere. This allowed ample sunlight to reach the surface of the earth which only enhanced the amount of photosynthesis done on earth. Moon light was also able to reach the earth's surface which initiated nocturnal life styles.

7) As the waters got oxygenated by the photosynthetic prokaryotes, eukaryotes started to emerge followed by soft bodied aquatic life. These precambrian life forms include goodfossilsof the firstmulti-celled animals like theediacaran biota,simplefossilsof a worm-liketrichophycus, the firstsponges,trilobitomorphs etc. There were also enigmatic soft-jellied creatures which were shaped like bags, disks, or quilts.

8) The cambrian period followed (543 mya)(Conway Morris, 2000),the most significant among all periods of earth's history. Because within a relatively short span of 5-10 million years, it produced the largest assortment of fossils that have been ever found. This cambrian explosion is believed to have given rise to most of the modern animal phyla, apart from the many other phyla which have gone extinct.

9)Biological evolution which started out in the water with small photosynthetic and non-photosynthetic life forms, continued on in the land giving rise to the non-aquatic (terrestrial) members of the plant and animal kingdom.

The evolution of the plant kingdom continued on land consecutively with the first green plants originating (mid ordovician period), followed by the first vascular plants (silurian), first clubmosses, ferns and horsetails, the first seed-bearing plants and the first trees (devonian), the first coal forests (pennsylvanian), first true gymnosperms & mosses (permian).

The evolution of the animal kingdom on land continued successively with first amphibians and first wingless insects (devonian), the first land vertebrates (mississippian), first reptiles

(pennsylvanian), first mammals & crocodiles (triassic), first birds and lizards (jurassic), first apes (miocene), first large mammals (paleocene).

10) During early evolution some free living bacteria were subjected to get into a symbiotic relationship with some eukaryotes by entering their cells and becoming one of their organelles, they are now known as mitochondria. This event initiated the evolution of the animal and plant kingdoms. There were some photosynthetic free living bacteria which entered these transformed eukaryotic cells and became an organelle now known as chloroplast, thereby giving rise to the evolution of the plant kingdom.

11) Humans are believed to be the most advanced among organisms. This is indicated by the large brain to body ratio seen in humans. In molecular terms humans have the largest collection of transcription factors amongst all sequenced animals. These evidences only show that the human body (especially the brain) is the most regulated or in other words the most complex among all organisms.

5. References:

- Abbott, M.K., and Kaufman, T.C. (1986). The relationship between the functional complexity and the molecular organization of the Antennapedia locus of *Drosophila melanogaster*. *Genetics* **114**, 919-942.
- Ackermann, M., and Chao, L. (2004). Evolution of cooperation: two for one? *Curr Biol* **14**, R73-74.
- Alberts, B. (1994). *Molecular biology of the cell*. Garland Pub.: New York.
- Alberts, B. (2002). *Molecular biology of the cell*. Garland Science: New York.
- Albrich, W.C., Monnet, D.L., and Harbarth, S. (2004). Antibiotic selection pressure and resistance in *Streptococcus pneumoniae* and *Streptococcus pyogenes*. *Emerg Infect Dis* **10**, 514-517.
- Allwood, A.C., Walter, M.R., Kamber, B.S., Marshall, C.P., and Burch, I.W. (2006). Stromatolite reef from the Early Archaean era of Australia. *Nature* **441**, 714-718.
- Artemenko, Y., Batsios, P., Borleis, J., Gagnon, Z., Lee, J., Rohlf, M., Sanseau, D., Willard, S.S., Schleicher, M., and Devreotes, P.N. (2012). Tumor suppressor Hippo/MST1 kinase mediates chemotaxis by regulating spreading and adhesion. *Proc Natl Acad Sci U S A* **109**, 13632-13637.
- Arthur, W. (1997). *The Origins of Animal Bodyplans*. Cambridge Press: London.
- Bagai, S., Puri, A., Blumenthal, R., and Sarkar, D.P. (1993). Hemagglutinin-neuraminidase enhances F protein-mediated membrane fusion of reconstituted Sendai virus envelopes with cells. *J Virol* **67**, 3312-3318.
- Banister, C.E. (2011). Review of Epigenetics: A Reference Manual: A book edited by Jeffrey M. Craig and Nicholas C. Wong. *Epigenetics* **7**, 963-964.
- Bao, Y., and Shen, X. (2011). SnapShot: Chromatin remodeling: INO80 and SWR1. *Cell* **144**, 158-158 e152.
- Bell, M.A.F., S. A. (1994). *The evolutionary biology of the threespine stickleback*. Oxford University Press: Oxford.
- Bennett, P.M. (2008). Plasmid encoded antibiotic resistance: acquisition and transfer of antibiotic resistance genes in bacteria. *Br J Pharmacol* **153 Suppl 1**, S347-357.
- Bignell, C., and Thomas, C.M. (2001). The bacterial ParA-ParB partitioning proteins. *J Biotechnol* **91**, 1-34.
- Blanc, G., and Wolfe, K.H. (2004). Functional divergence of duplicated genes formed by polyploidy during Arabidopsis evolution. *Plant Cell* **16**, 1679-1691.
- Blanton, M.J.G.R.L. SEM of Dictyostelium developmental stages.: Biological Sciences Electron Microscopy Laboratory, Texas Tech University.

- Bolourani, P., Spiegelman, G.B., and Weeks, G. (2006). Delineation of the roles played by RasG and RasC in cAMP-dependent signal transduction during the early development of *Dictyostelium discoideum*. *Mol Biol Cell* 17, 4543-4550.
- Bonner, J.T. (2000). *First signals: the evolution of multicellular development*. Princeton University Press: Princeton.
- Bork, P., Sander, C., and Valencia, A. (1992). An ATPase domain common to prokaryotic cell cycle proteins, sugar kinases, actin, and hsp70 heat shock proteins. *Proc Natl Acad Sci U S A* 89, 7290-7294.
- Bowring, S.A., Grotzinger, J.P., Isachsen, C.E., Knoll, A.H., Pelechaty, S.M., and Kolosov, P. (1993). Calibrating rates of early Cambrian evolution. *Science* 261, 1293-1298.
- Boyer, L.A., and Peterson, C.L. (2000). Actin-related proteins (Arps): conformational switches for chromatin-remodeling machines? *Bioessays* 22, 666-672.
- Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal Biochem* 72, 248-254.
- Brock, D.A., Douglas, T.E., Queller, D.C., and Strassmann, J.E. (2011). Primitive agriculture in a social amoeba. *Nature* 469, 393-396.
- Brown, K.S. (1999). Deep Green rewrites evolutionary history of plants. *Science* 285, 990-991.
- Burgess, K.S., Etterson, J.R., and Galloway, L.F. (2007). Artificial selection shifts flowering phenology and other correlated traits in an autotetraploid herb. *Heredity (Edinb)* 99, 641-648.
- Butlin, R.K., Galindo, J., and Grahame, J.W. (2008). Review. Sympatric, parapatric or allopatric: the most important way to classify speciation? *Philos Trans R Soc Lond B Biol Sci* 363, 2997-3007.
- Butterfield, N.J. (2009). Oxygen, animals and oceanic ventilation: an alternative view. *Geobiology* 7, 1-7.
- Capelson, M., Doucet, C., and Hetzer, M.W. (2011). Nuclear pore complexes: guardians of the nuclear genome. *Cold Spring Harb Symp Quant Biol* 75, 585-597.
- Carballido-Lopez, R. (2006). The bacterial actin-like cytoskeleton. *Microbiol Mol Biol Rev* 70, 888-909.
- Carrier, M.F. (1990). Actin polymerization and ATP hydrolysis. *Adv Biophys* 26, 51-73.
- Casadesus, J., and Low, D. (2006). Epigenetic gene regulation in the bacterial world. *Microbiol Mol Biol Rev* 70, 830-856.
- Celniker, S.E., and Rubin, G.M. (2003). The *Drosophila melanogaster* genome. *Annu Rev Genomics Hum Genet* 4, 89-117.
- Chahwan, R., Wontakal, S.N., and Roa, S. (2010). Crosstalk between genetic and epigenetic information through cytosine deamination. *Trends Genet* 26, 443-448.
- Chahwan, R., Wontakal, S.N., and Roa, S. (2011). The multidimensional nature of epigenetic information and its role in disease. *Discov Med* 11, 233-243.

- Chen, L., Cai, Y., Jin, J., Florens, L., Swanson, S.K., Washburn, M.P., Conaway, J.W., and Conaway, R.C. (2011). Subunit organization of the human INO80 chromatin remodeling complex: an evolutionarily conserved core complex catalyzes ATP-dependent nucleosome remodeling. *J Biol Chem* 286, 11283-11289.
- Chen, R., Quinones-Mateu, M.E., and Mansky, L.M. (2004). Drug resistance, virus fitness and HIV-1 mutagenesis. *Curr Pharm Des* 10, 4065-4070.
- Cherry, J.L. (2010). Highly expressed and slowly evolving proteins share compositional properties with thermophilic proteins. *Mol Biol Evol* 27, 735-741.
- Chouard, T. (2010). Evolution: Revenge of the hopeful monster. *Nature* 463, 864-867.
- Cohen, K.M., Finney, S., and Gibbard, P.L. (2012). International Chronostratigraphic Chart. Chart reproduced for the 34th International Geological Congress, Brisbane, Australia.
- Cokol, M., Nair, R., and Rost, B. (2000). Finding nuclear localization signals. *EMBO Rep* 1, 411-415.
- Consortium, I.C.G.S. (2004). Sequence and comparative analysis of the chicken genome provide unique perspectives on vertebrate evolution. *Nature* 432, 695-716.
- Conway Morris, S. (2000). The Cambrian "explosion": slow-fuse or megatonnage? *Proc Natl Acad Sci U S A* 97, 4426-4429.
- Costa, C.F., Rommelaere, H., Waterschoot, D., Sethi, K.K., Nowak, K.J., Laing, N.G., Ampe, C., and Machesky, L.M. (2004). Myopathy mutations in alpha-skeletal-muscle actin cause a range of molecular defects. *J Cell Sci* 117, 3367-3377.
- Crespi, B.J. (2001). The evolution of social behavior in microorganisms. *Trends Ecol Evol* 16, 178-183.
- Crooks, G.E., Hon, G., Chandonia, J.M., and Brenner, S.E. (2004). WebLogo: a sequence logo generator. *Genome Res* 14, 1188-1190.
- Csete, M.E., and Doyle, J.C. (2002). Reverse engineering of biological complexity. *Science* 295, 1664-1669.
- Cui, L., Wall, P.K., Leebens-Mack, J.H., Lindsay, B.G., Soltis, D.E., Doyle, J.J., Soltis, P.S., Carlson, J.E., Arumuganathan, K., Barakat, A., Albert, V.A., Ma, H., and dePamphilis, C.W. (2006). Widespread genome duplications throughout the history of flowering plants. *Genome Res* 16, 738-749.
- D.G. Shu, H.L.L., S. Conway Morris, X.L. Zhang, S.X. Hu, L. Chen, J. Han, M. Zhu, Y. Li, and, and Chen, L.Z. (1999). Lower Cambrian vertebrates from south China. *Nature*.
- Dalton, R. (2010). Ancient DNA set to rewrite human history. *Nature* 465, 148-149.
- Dayel, M.J., and Mullins, R.D. (2004). Activation of Arp2/3 complex: addition of the first subunit of the new filament by a WASP protein triggers rapid ATP hydrolysis on Arp2. *PLoS Biol* 2, E91.
- de Rosa, R., Grenier, J.K., Andreeva, T., Cook, C.E., Adoutte, A., Akam, M., Carroll, S.B., and Balavoine, G. (1999). Hox genes in brachiopods and priapulids and protostome evolution. *Nature* 399, 772-776.
- Deal, R.B., and Henikoff, S. (2010). Gene regulation: A chromatin thermostat. *Nature* 463, 887-888.

- Dodd, D.M.B. (1989). Reproductive Isolation as a Consequence of Adaptive Divergence in *Drosophila pseudoobscura*. *Evolution* Vol. 43, 1308-1311.
- Doebley, J., Stec, A., and Hubbard, L. (1997). The evolution of apical dominance in maize. *Nature* 386, 485-488.
- Dominguez, R. (2004). Actin-binding proteins--a unifying hypothesis. *Trends Biochem Sci* 29, 572-578.
- Doolittle, R.F., and York, A.L. (2002). Bacterial actins? An evolutionary perspective. *Bioessays* 24, 293-296.
- Drummond, D.A., Bloom, J.D., Adami, C., Wilke, C.O., and Arnold, F.H. (2005). Why highly expressed proteins evolve slowly. *Proc Natl Acad Sci U S A* 102, 14338-14343.
- Duret, L., and Mouchiroud, D. (2000). Determinants of substitution rates in mammalian genes: expression pattern affects selection intensity but not mutation rate. *Mol Biol Evol* 17, 68-74.
- Dynes, J.L., Clark, A.M., Shaulsky, G., Kuspa, A., Loomis, W.F., and Firtel, R.A. (1994). LagC is required for cell-cell interactions that are essential for cell-type differentiation in *Dictyostelium*. *Genes Dev* 8, 948-958.
- Eichinger, L., and Noegel, A.A. (2003). Crawling into a new era-the *Dictyostelium* genome project. *Embo J* 22, 1941-1946.
- Eichinger, L., Pachebat, J.A., Glockner, G., Rajandream, M.A., Sucgang, R., Berriman, M., Song, J., Olsen, R., Szafranski, K., Xu, Q., Tunggal, B., Kummerfeld, S., Madera, M., Konfortov, B.A., Rivero, F., Bankier, A.T., Lehmann, R., Hamlin, N., Davies, R., Gaudet, P., Fey, P., Pilcher, K., Chen, G., Saunders, D., Sodergren, E., Davis, P., Kerhornou, A., Nie, X., Hall, N., Anjard, C., Hemphill, L., Bason, N., Farbrother, P., Desany, B., Just, E., Morio, T., Rost, R., Churcher, C., Cooper, J., Haydock, S., van Driessche, N., Cronin, A., Goodhead, I., Muzny, D., Mourier, T., Pain, A., Lu, M., Harper, D., Lindsay, R., Hauser, H., James, K., Quiles, M., Madan Babu, M., Saito, T., Buchrieser, C., Wardroper, A., Felder, M., Thangavelu, M., Johnson, D., Knights, A., Loulseged, H., Mungall, K., Oliver, K., Price, C., Quail, M.A., Urushihara, H., Hernandez, J., Rabinowitsch, E., Steffen, D., Sanders, M., Ma, J., Kohara, Y., Sharp, S., Simmonds, M., Spiegler, S., Tivey, A., Sugano, S., White, B., Walker, D., Woodward, J., Winckler, T., Tanaka, Y., Shaulsky, G., Schleicher, M., Weinstock, G., Rosenthal, A., Cox, E.C., Chisholm, R.L., Gibbs, R., Loomis, W.F., Platzer, M., Kay, R.R., Williams, J., Dear, P.H., Noegel, A.A., Barrell, B., and Kuspa, A. (2005). The genome of the social amoeba *Dictyostelium discoideum*. *Nature* 435, 43-57.
- Erickson, H.P. (2007). Evolution of the cytoskeleton. *Bioessays* 29, 668-677.
- Esue, O., Rupprecht, L., Sun, S.X., and Wirtz, D. (Jan. 2010). Dynamics of the bacterial intermediate filament crescentin in vitro and in vivo. *PLoS One* 5, e8855.
- Evans, J.P., and Florman, H.M. (2002). The state of the union: the cell biology of fertilization. *Nat Cell Biol* 4 Suppl, s57-63.
- Faguy, D.M., and Jarrell, K.F. (1999). A twisted tale: the origin and evolution of motility and chemotaxis in prokaryotes. *Microbiology* 145 (Pt 2), 279-281.

Faix, J., Kreppel, L., Shaulsky, G., Schleicher, M., and Kimmel, A.R. (2004). A rapid and efficient method to generate multiple gene disruptions in *Dictyostelium discoideum* using a single selectable marker and the Cre-loxP system. *Nucleic Acids Res* 32, e143.

Feng, W., Benko, A.L., Lee, J.H., Stanford, D.R., and Hopper, A.K. (1999). Antagonistic effects of NES and NLS motifs determine *S. cerevisiae* Rna1p subcellular distribution. *J Cell Sci* 112 (Pt 3), 339-347.

Fiegna, F., and Velicer, G.J. (2003). Competitive fates of bacterial social parasites: persistence and self-induced extinction of *Myxococcus xanthus* cheaters. *Proc Biol Sci* 270, 1527-1534.

Figge, R.M., Divakaruni, A.V., and Gober, J.W. (2004). MreB, the cell shape-determining bacterial actin homologue, co-ordinates cell wall morphogenesis in *Caulobacter crescentus*. *Mol Microbiol* 51, 1321-1332.

Fike, D.A., Grotzinger, J.P., Pratt, L.M., and Summons, R.E. (2006). Oxidation of the Ediacaran ocean. *Nature* 444, 744-747.

Finn, R.D., Mistry, J., Schuster-Bockler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., Moxon, S., Marshall, M., Khanna, A., Durbin, R., Eddy, S.R., Sonnhammer, E.L., and Bateman, A. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res* 34, D247-251.

Flaherty, K.M., McKay, D.B., Kabsch, W., and Holmes, K.C. (1991). Similarity of the three-dimensional structures of actin and the ATPase fragment of a 70-kDa heat shock cognate protein. *Proc Natl Acad Sci U S A* 88, 5041-5045.

Fostinis, Y., Theodoropoulos, P.A., Gravanis, A., and Stournaras, C. (1992). Heat shock protein HSP90 and its association with the cytoskeleton: a morphological study. *Biochem Cell Biol* 70, 779-786.

Frankel, S., and Mooseker, M.S. (1996). The actin-related proteins. *Curr Opin Cell Biol* 8, 30-37.

Gancedo, C., and Flores, C.L. (2008). Moonlighting proteins in yeasts. *Microbiol Mol Biol Rev* 72, 197-210, table of contents.

Gould, S.J., & Eldredge, Niles. (1977). Punctuated equilibria: the tempo and mode of evolution reconsidered. *Paleobiology* 3 (2), p.145.

Gould, S.J., and Eldredge, N. (1993). Punctuated equilibrium comes of age. *Nature* 366, 223-227.

Grant P. R., G.B.R. (1994). Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution*, 297-316.

Grant, P.R., and Grant, B.R. (2002). Unpredictable evolution in a 30-year study of Darwin's finches. *Science* 296, 707-711.

Grava, S., Dumoulin, P., Madania, A., Tarasov, I., and Winsor, B. (2000). Functional analysis of six genes from chromosomes XIV and XV of *Saccharomyces cerevisiae* reveals YOR145c as an essential gene and YNL059c/ARP5 as a strain-dependent essential gene encoding nuclear proteins. *Yeast* 16, 1025-1033.

Grenier, J.K., Garber, T.L., Warren, R., Whittington, P.M., and Carroll, S. (1997). Evolution of the entire arthropod Hox gene set predated the origin and radiation of the onychophoran/arthropod clade. *Curr Biol* 7, 547-553.

- Grotzinger, J.P., and Knoll, A.H. (1999). Stromatolites in Precambrian carbonates: evolutionary mileposts or environmental dipsticks? *Annu Rev Earth Planet Sci* 27, 313-358.
- Hamilton, B.S., Whittaker, G.R., and Daniel, S. (2012). Influenza virus-mediated membrane fusion: determinants of hemagglutinin fusogenic activity and experimental approaches for assessing virus fusion. *Viruses* 4, 1144-1168.
- Hart, A.G., Stafford, R., Smith, A.L., and Goodenough, A.E. (2010). Evidence for contemporary evolution during Darwin's lifetime. *Curr Biol* 20, R95.
- Hastings, K.E. (1996). Strong evolutionary conservation of broadly expressed protein isoforms in the troponin I gene family and other vertebrate gene families. *J Mol Evol* 42, 631-640.
- Havey, J.C., Vecchiarelli, A.G., and Funnell, B.E. (2012 Jan). ATP-regulated interactions between P1 ParA, ParB and non-specific DNA that are stabilized by the plasmid partition site, parS. *Nucleic Acids Res* 40, 801-812.
- Hennig, W. (2004). The revolution of the biology of the genome. *Cell Res* 14, 1-7.
- Herrmann, H., and Strelkov, S.V. (2011). History and phylogeny of intermediate filaments: now in insects. *BMC Biol* 9, 16.
- Hirschman, J.E., Balakrishnan, R., Christie, K.R., Costanzo, M.C., Dwight, S.S., Engel, S.R., Fisk, D.G., Hong, E.L., Livstone, M.S., Nash, R., Park, J., Oughtred, R., Skrzypek, M., Starr, B., Theesfeld, C.L., Williams, J., Andrada, R., Binkley, G., Dong, Q., Lane, C., Miyasato, S., Sethuraman, A., Schroeder, M., Thanawala, M.K., Weng, S., Dolinski, K., Botstein, D., and Cherry, J.M. (2006). Genome Snapshot: a new resource at the Saccharomyces Genome Database (SGD) presenting an overview of the *Saccharomyces cerevisiae* genome. *Nucleic Acids Res* 34, D442-445.
- Holmes, D.S., and Quigley, M. (1981). A rapid boiling method for the preparation of bacterial plasmids. *Anal Biochem* 114, 193-197.
- Holmes, E.C. (2007). Viral evolution in the genomic age. *PLoS Biol* 5, e278.
- Hoskin, C.J., Higgie, M., McDonald, K.R., and Moritz, C. (2005). Reinforcement drives rapid allopatric speciation. *Nature* 437, 1353-1356.
- Howard, P.K., Sefton, B.M., and Firtel, R.A. (1993). Tyrosine phosphorylation of actin in Dictyostelium associated with cell-shape changes. *Science* 259, 241-244.
- Hsie, A.W., Recio, L., Katz, D.S., Lee, C.Q., Wagner, M., and Schenley, R.L. (1986). Evidence for reactive oxygen species inducing mutations in mammalian cells. *Proc Natl Acad Sci U S A* 83, 9616-9620.
- Huber, S.K., De Leon, L.F., Hendry, A.P., Bermingham, E., and Podos, J. (2007). Reproductive isolation of sympatric morphs in a population of Darwin's finches. *Proc Biol Sci* 274, 1709-1714.
- Huberts, D.H., and van der Klei, I.J. (2010). Moonlighting proteins: an intriguing mode of multitasking. *Biochim Biophys Acta* 1803, 520-525.
- Iijima, M., and Devreotes, P. (2002). Tumor suppressor PTEN mediates sensing of chemoattractant gradients. *Cell* 109, 599-610.

- Iyer, L.M., Balaji, S., Koonin, E.V., and Aravind, L. (2006). Evolutionary genomics of nucleo-cytoplasmic large DNA viruses. *Virus Res* 117, 156-184.
- Jackson, J.B., and Cheetham, A.H. (1990). Evolutionary significance of morphospecies: a test with cheilostome bryozoa. *Science* 248, 579-583.
- Jeffery, C.J. (1999). Moonlighting proteins. *Trends Biochem Sci* 24, 8-11.
- Jeffery, C.J. (2003). Moonlighting proteins: old proteins learning new tricks. *Trends Genet* 19, 415-417.
- Jeffery, C.J. (2004). Molecular mechanisms for multitasking: recent crystal structures of moonlighting proteins. *Curr Opin Struct Biol* 14, 663-668.
- Jin, J., Cai, Y., Yao, T., Gottschalk, A.J., Florens, L., Swanson, S.K., Gutierrez, J.L., Coleman, M.K., Workman, J.L., Mushegian, A., Washburn, M.P., Conaway, R.C., and Conaway, J.W. (2005). A mammalian chromatin remodeling complex with similarities to the yeast INO80 complex. *J Biol Chem* 280, 41207-41212.
- Jordan, I.K., Rogozin, I.B., Wolf, Y.I., and Koonin, E.V. (2002). Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. *Genome Res* 12, 962-968.
- Joseph, J.M., Fey, P., Ramalingam, N., Liu, X.I., Rohlf, M., Noegel, A.A., Muller-Taubenberger, A., Glockner, G., and Schleicher, M. (2008). The actinome of Dictyostelium discoideum in comparison to actins and actin-related proteins from other organisms. *PLoS One* 3, e2654.
- Jungbluth, A., Eckerskorn, C., Gerisch, G., Lottspeich, F., Stocker, S., and Schweiger, A. (1995). Stress-induced tyrosine phosphorylation of actin in Dictyostelium cells and localization of the phosphorylation site to tyrosine-53 adjacent to the DNase I binding loop. *FEBS Lett* 375, 87-90.
- Jungbluth, A., von Arnim, V., Biegelmann, E., Humbel, B., Schweiger, A., and Gerisch, G. (1994). Strong increase in the tyrosine phosphorylation of actin upon inhibition of oxidative phosphorylation: correlation with reversible rearrangements in the actin skeleton of Dictyostelium cells. *J Cell Sci* 107 (Pt 1), 117-125.
- Jurka, J., Bao, W., and Kojima, K.K. (2011). Families of transposable elements, population structure and the origin of species. *Biol Direct* 6, 44.
- Kabsch, W., and Holmes, K.C. (1995). The actin fold. *Faseb J* 9, 167-174.
- Kabsch, W., Mannherz, H.G., Suck, D., Pai, E.F., and Holmes, K.C. (1990). Atomic structure of the actin:DNase I complex. *Nature* 347, 37-44.
- Kaiser, D. (2001). Building a multicellular organism. *Annu Rev Genet* 35, 103-123.
- Kakunaga, T., Leavitt, J., and Hamada, H. (1984). A mutation in actin associated with neoplastic transformation. *Fed Proc* 43, 2275-2279.
- Kamei, N., and Glabe, C.G. (2003). The species-specific egg receptor for sea urchin sperm adhesion is EBR1, a novel ADAMTS protein. *Genes Dev* 17, 2502-2507.
- Kerr, R.A. (1993). Evolution's Big Bang Gets Even More Explosive. *Science* 261, 1274-1275.

- Kishi, Y., Clements, C., Mahadeo, D.C., Cotter, D.A., and Sameshima, M. (1998). High levels of actin tyrosine phosphorylation: correlation with the dormant state of *Dictyostelium* spores. *J Cell Sci* **111** (Pt 19), 2923-2932.
- Kishi, Y., Mahadeo, D., Cervi, D.N., Clements, C., Cotter, D.A., and Sameshima, M. (2000). Glucose-induced pathways for actin tyrosine dephosphorylation during *Dictyostelium* spore germination. *Exp Cell Res* **261**, 187-198.
- Klein, C. (2005). Some Precambrian banded iron-formations (BIFs) from around the world: Their age, geologic setting, mineralogy, metamorphism, geochemistry, and origins. *American Mineralogist* **90**, p. 1473-1499.
- Komeili, A., Li, Z., Newman, D.K., and Jensen, G.J. (2006). Magnetosomes are cell membrane invaginations organized by the actin-like protein MamK. *Science* **311**, 242-245.
- Kreis, T., and Vale, R. (1999). Guidebook to the cytoskeletal and motor proteins. Oxford University Press: Oxford; New York.
- Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol* **305**, 567-580.
- Kruse, T., Bork-Jensen, J., and Gerdes, K. (2005). The morphogenetic MreBCD proteins of *Escherichia coli* form an essential membrane-bound complex. *Mol Microbiol* **55**, 78-89.
- Kruse, T., and Gerdes, K. (2005). Bacterial DNA segregation by the actin-like MreB protein. *Trends Cell Biol* **15**, 343-345.
- Kuma, K., Iwabe, N., and Miyata, T. (1995). Functional constraints against variations on molecules from the tissue level: slowly evolving brain-specific genes demonstrated by protein kinase and immunoglobulin supergene families. *Mol Biol Evol* **12**, 123-130.
- la Cour, T., Kiemer, L., Molgaard, A., Gupta, R., Skriver, K., and Brunak, S. (2004). Analysis and prediction of leucine-rich nuclear export signals. *Protein Eng Des Sel* **17**, 527-536.
- Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* **227**, 680-685.
- Lenski, R.E. (2004). Microbial evolution in action. *Microbiology Today*, 158.
- Lenski, R.E. (2006). *E. coli* as a model organism in evolutionary genetics. Oxford University Press: New York.
- Levinton, J.S. (1992). The big bang of animal evolution. *Sci Am* **267**, 84-91.
- Levinton, J.S. (2008). The Cambrian Explosion: How Do We Use the Evidence. *BioScience Vol.* **58**, 855-864.
- Liku, M.E., Nguyen, V.Q., Rosales, A.W., Irie, K., and Li, J.J. (2005). CDK phosphorylation of a novel NLS-NES module distributed between two subunits of the Mcm2-7 complex prevents chromosomal rereplication. *Mol Biol Cell* **16**, 5026-5039.

- Linville, B.J., Stewart, J.R., Ecay, T.W., Herbert, J.F., Parker, S.L., and Thompson, M.B. (2010). Placental calcium provision in a lizard with prolonged oviductal egg retention. *J Comp Physiol B* 180, 221-227.
- Liu, X., Shu, S., Hong, M.S., Levine, R.L., and Korn, E.D. (2006). Phosphorylation of actin Tyr-53 inhibits filament nucleation and elongation and destabilizes filaments. *Proc Natl Acad Sci U S A* 103, 13694-13699.
- Lloyd, C. (2011). Dynamic microtubules and the texture of plant cell walls. *Int Rev Cell Mol Biol* 287, 287-329.
- Loftus, B., Anderson, I., Davies, R., Alsmark, U.C., Samuelson, J., Amedeo, P., Roncaglia, P., Berriman, M., Hirt, R.P., Mann, B.J., Nozaki, T., Suh, B., Pop, M., Duchene, M., Ackers, J., Tannich, E., Leippe, M., Hofer, M., Bruchhaus, I., Willhoeft, U., Bhattacharya, A., Chillingworth, T., Churcher, C., Hance, Z., Harris, B., Harris, D., Jagels, K., Moule, S., Mungall, K., Ormond, D., Squares, R., Whitehead, S., Quail, M.A., Rabinowitsch, E., Norbertczak, H., Price, C., Wang, Z., Guillen, N., Gilchrist, C., Stroup, S.E., Bhattacharya, S., Lohia, A., Foster, P.G., Sicheritz-Ponten, T., Weber, C., Singh, U., Mukherjee, C., El-Sayed, N.M., Petri, W.A., Jr., Clark, C.G., Embley, T.M., Barrell, B., Fraser, C.M., and Hall, N. (2005). The genome of the protist parasite *Entamoeba histolytica*. *Nature* 433, 865-868.
- Logsdon, J.M., Jr., and Doolittle, W.F. (1997). Origin of antifreeze protein genes: a cool tale in molecular evolution. *Proc Natl Acad Sci U S A* 94, 3485-3487.
- Losos, J.B., Schoener, T.W., Warheit, K.I., and Creer, D. (2001). Experimental studies of adaptive differentiation in Bahamian *Anolis* lizards. *Genetica* 112-113, 399-415.
- Lutkenhaus, J. (2012 Sep). The ParA/MinD family puts things in their place. *Trends Microbiol* 20, 411-418.
- Mackay, T.F., Stone, E.A., and Ayroles, J.F. (2009). The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10, 565-577.
- MacWilliams HK, B.J. (1979). The prestalk-prespore pattern in cellular slime molds. *Differentiation*.
- Malchow, D., Nagele, B., Schwarz, H., and Gerisch, G. (1972). Membrane-bound cyclic AMP phosphodiesterase in chemotactically responding cells of *Dictyostelium discoideum*. *Eur J Biochem* 28, 136-142.
- Mallarino, R., Grant, P.R., Grant, B.R., Herrel, A., Kuo, W.P., and Abzhanov, A. (2011). Two developmental modules establish 3D beak-shape variation in Darwin's finches. *Proc Natl Acad Sci U S A* 108, 4057-4062.
- Margolin, W. (2004). Bacterial shape: concave coiled coils curve caulobacter. *Curr Biol* 14, R242-244.
- Marie Curie SPECIATION Network, B.R., DeBelle A, Kerth C, Snook RR, Beukeboom LW, Castillo Cajas RF, Diao W, Maan ME, Paolucci S, Weissing FJ, van de Zande L, Hoikkala A, Geuverink E, Jennings J, Kankare M, Knott KE, Tyukmaeva VI, Zoumadakis C, Ritchie MG, Barker D, Immonen E, Kirkpatrick M, Noor M, Macias Garcia C, Schmitt T, Schilthuizen M. (2012). What do we need to know about speciation? *Trends Ecol Evol*.
- Mattila, T.M., and Bokma, F. (2008). Extant mammal body masses suggest punctuated equilibrium. *Proc Biol Sci* 275, 2195-2199.

- McKinney, E.C., and Meagher, R.B. (1998). Members of the Arabidopsis actin gene family are widely dispersed in the genome. *Genetics* **149**, 663-675.
- Meagher, R.B., Kandasamy, M.K., McKinney, E.C., and Roy, E. (2009). Chapter 5. Nuclear actin-related proteins in epigenetic control. *Int Rev Cell Mol Biol* **277**, 157-215.
- Miller, B.M., and Trybus, K.M. (2008). Functional effects of nemaline myopathy mutations on human skeletal alpha-actin. *J Biol Chem* **283**, 19379-19388.
- Miller, R.L. (1985). *Sperm chemo-orientation in the metazoa*. Academic Press: New York.
- Mingorance, J., Rivas, G., Velez, M., Gomez-Puertas, P., and Vicente, M. (Aug. 2011). Strong FtsZ is with the force: mechanisms to constrict bacteria. *Trends Microbiol* **18**, 348-356.
- Moller-Jensen, J., Borch, J., Dam, M., Jensen, R.B., Roepstorff, P., and Gerdes, K. (2003). Bacterial mitosis: ParM of plasmid R1 moves plasmid DNA by an actin-like insertional polymerization mechanism. *Mol Cell* **12**, 1477-1487.
- Monastersky, R. (1993). Siberian Rocks Clock Biological Big Bang. *Science News*.
- Mulkidjanian, A.Y., Bychkov, A.Y., Dibrova, D.V., Galperin, M.Y., and Koonin, E.V. (2012). Origin of first cells at terrestrial, anoxic geothermal fields. *Proc Natl Acad Sci U S A* **109**, E821-830.
- Muller, J., Oma, Y., Vallar, L., Friederich, E., Poch, O., and Winsor, B. (2005). Sequence and comparative genomic analysis of actin-related proteins. *Mol Biol Cell* **16**, 5736-5748.
- Murthy, K., and Wadsworth, P. (2005). Myosin-II-dependent localization and dynamics of F-actin during cytokinesis. *Curr Biol* **15**, 724-731.
- Nash, J.M. (1995). When Life Exploded. In: *Time*, pp. 67,68,70,74.
- Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997). A neural network method for identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Int J Neural Syst* **8**, 581-599.
- Noguchi, T.Q., Kanzaki, N., Ueno, H., Hirose, K., and Uyeda, T.Q. (2007). A novel system for expressing toxic actin mutants in Dictyostelium and purification and characterization of a dominant lethal yeast actin mutant. *J Biol Chem* **282**, 27721-27727.
- Olson, T.M., Michels, V.V., Thibodeau, S.N., Tai, Y.S., and Keating, M.T. (1998). Actin mutations in dilated cardiomyopathy, a heritable form of heart failure. *Science* **280**, 750-752.
- Otterbein, L.R., Graceffa, P., and Dominguez, R. (2001). The crystal structure of uncomplexed actin in the ADP state. *Science* **293**, 708-711.
- Page, R.B., Boley, M.A., Smith, J.J., Putta, S., and Voss, S.R. (2010). Microarray analysis of a salamander hopeful monster reveals transcriptional signatures of paedomorphic brain development. *BMC Evol Biol* **10**, 199.
- Palmer, J. (2011). Amoebas show primitive farming behaviour as they travel.

Paran, I., and Zamir, D. (2003). Quantitative traits in plants: beyond the QTL. *Trends Genet* 19, 303-306.

Paredez, A.R., Somerville, C.R., and Ehrhardt, D.W. (2006). Visualization of cellulose synthase demonstrates functional association with microtubules. *Science* 312, 1491-1495.

Pena-Castillo, L., Tasan, M., Myers, C.L., Lee, H., Joshi, T., Zhang, C., Guan, Y., Leone, M., Pagnani, A., Kim, W.K., Krumpelman, C., Tian, W., Obozinski, G., Qi, Y., Mostafavi, S., Lin, G.N., Berriz, G.F., Gibbons, F.D., Lanckriet, G., Qiu, J., Grant, C., Barutcuoglu, Z., Hill, D.P., Warde-Farley, D., Grouios, C., Ray, D., Blake, J.A., Deng, M., Jordan, M.I., Noble, W.S., Morris, Q., Klein-Seetharaman, J., Bar-Joseph, Z., Chen, T., Sun, F., Troyanskaya, O.G., Marcotte, E.M., Xu, D., Hughes, T.R., and Roth, F.P. (2008). A critical assessment of *Mus musculus* gene function prediction using integrated genomic evidence. *Genome Biol* 9 Suppl 1, S2.

Pennisi, E. (2001). Molecular evolution. Genome duplications: the stuff of evolution? *Science* 294, 2458-2460.

Petren, K., Grant, P.R., Grant, B.R., and Keller, L.F. (2005). Comparative landscape genetics and the adaptive radiation of Darwin's finches: the role of peripheral isolation. *Mol Ecol* 14, 2943-2957.

Plomin, R., Haworth, C.M., and Davis, O.S. (2009). Common disorders are quantitative traits. *Nat Rev Genet* 10, 872-878.

Plotch, S.J., O'Hara, B., Morin, J., Palant, O., LaRocque, J., Bloom, J.D., Lang, S.A., Jr., DiGrandi, M.J., Bradley, M., Nilakantan, R., and Gluzman, Y. (1999). Inhibition of influenza A virus replication by compounds interfering with the fusogenic function of the viral hemagglutinin. *J Virol* 73, 140-151.

Poch, O., and Winsor, B. (1997). Who's who among the *Saccharomyces cerevisiae* actin-related proteins? A classification and nomenclature proposal for a large family. *Yeast* 13, 1053-1058.

Popp, D., and Robinson, R.C. (2011). Many ways to build an actin filament. *Mol Microbiol* 80, 300-308.

Procaccio, V., Salazar, G., Ono, S., Styers, M.L., Gearing, M., Davila, A., Jimenez, R., Juncos, J., Gutekunst, C.A., Meroni, G., Fontanella, B., Sontag, E., Sontag, J.M., Faundez, V., and Wainer, B.H. (2006). A mutation of beta -actin that alters depolymerization dynamics is associated with autosomal dominant developmental malformations, deafness, and dystonia. *Am J Hum Genet* 78, 947-960.

Queitsch, C., Sangster, T.A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature* 417, 618-624.

Rajesh, A.a.S., M. (2006). The actin cytoskeleton. Springer Verlag.

Randazzo, F.M., Seeger, M.A., Huss, C.A., Sweeney, M.A., Cecil, J.K., and Kaufman, T.C. (1993). Structural changes in the antennapedia complex of *Drosophila pseudoobscura*. *Genetics* 134, 319-330.

Rechsteiner, M., and Rogers, S.W. (1996). PEST sequences and regulation by proteolysis. *Trends Biochem Sci* 21, 267-271.

Robert, F., and Chaussidon, M. (2006). A palaeotemperature curve for the Precambrian oceans based on silicon isotopes in cherts. *Nature* 443, 969-972.

Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* 234, 364-368.

Ronschaugen, M., McGinnis, N., and McGinnis, W. (2002). Hox protein mutation and macroevolution of the insect body plan. *Nature* 415, 914-917.

Rubenstein, P., and Deuchler, J. (1979). Acetylated and nonacetylated actins in Dictyostelium discoideum. *J Biol Chem* 254, 11142-11147.

Rubenstein, P., Smith, P., Deuchler, J., and Redman, K. (1981). NH₂-terminal acetylation of Dictyostelium discoideum actin in a cell-free protein-synthesizing system. *J Biol Chem* 256, 8149-8155.

Rubenstein, P.A. (1990). The functional importance of multiple actin isoforms. *Bioessays* 12, 309-315.

Rubin-Pitel, S.B., and Zhao, H. (2006). Recent advances in biocatalysis by directed enzyme evolution. *Comb Chem High Throughput Screen* 9, 247-257.

Rutherford, S.L., and Lindquist, S. (1998). Hsp90 as a capacitor for morphological evolution. *Nature* 396, 336-342.

Saltzman, M.R., Young, S.A., Kump, L.R., Gill, B.C., Lyons, T.W., and Runnegar, B. (2011). Pulse of atmospheric oxygen during the late Cambrian. *Proc Natl Acad Sci U S A* 108, 3876-3881.

Sambrook, J., and Russell, D.W. (2001). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press: Cold Spring Harbor, N.Y.

Sameshima, M., Chijiwa, Y., Kishi, Y., and Hashimoto, Y. (1994). Novel actin rods appeared in spores of Dictyostelium discoideum. *Cell Struct Funct* 19, 189-194.

Sameshima, M., Kishi, Y., Osumi, M., Mahadeo, D., and Cotter, D.A. (2000). Novel actin cytoskeleton: actin tubules. *Cell Struct Funct* 25, 291-295.

Sameshima, M., Kishi, Y., Osumi, M., Mahadeo, D., and Cotter, D.A. (2002). Electron microscopy of actin rods and bundles in Dictyostelium discoideum by high-pressure freezing. *J Electron Microsc* (Tokyo) 51, 337-340.

Sameshima, M., Kishi, Y., Osumi, M., Minamikawa-Tachino, R., Mahadeo, D., and Cotter, D.A. (2001). The formation of actin rods composed of actin tubules in Dictyostelium discoideum spores. *J Struct Biol* 136, 7-19.

Schlieper, D., Oliva, M.A., Andreu, J.M., and Lowe, J. (2005). Structure of bacterial tubulin BtubA/B: evidence for horizontal gene transfer. *Proc Natl Acad Sci U S A* 102, 9170-9175.

Schweiger, A., Mihalache, O., Ecke, M., and Gerisch, G. (1992). Stage-specific tyrosine phosphorylation of actin in Dictyostelium discoideum cells. *J Cell Sci* 102 (Pt 3), 601-609.

Seoighe, C., and Gehring, C. (2004). Genome duplication led to highly selective expansion of the Arabidopsis thaliana proteome. *Trends Genet* 20, 461-464.

Shen, B., and Lutkenhaus, J. (Nov. 2010). Differences in MinC/MinD sensitivity between polar and internal Z rings in Escherichia coli. *J Bacteriol* 193, 367-376.

- Shen, X., Ranallo, R., Choi, E., and Wu, C. (2003). Involvement of actin-related proteins in ATP-dependent chromatin remodeling. *Mol Cell* 12, 147-155.
- Sherr, C.J. (2004). Principles of tumor suppression. *Cell* 116, 235-246.
- Shih, Y.L., and Rothfield, L. (2006). The bacterial cytoskeleton. *Microbiol Mol Biol Rev* 70, 729-754.
- Shiomi, D., Yoshimoto, M., Homma, M., and Kawagishi, I. (2006). Helical distribution of the bacterial chemoreceptor via colocalization with the Sec protein translocation machinery. *Mol Microbiol* 60, 894-906.
- Simonsen, L., Viboud, C., Grenfell, B.T., Dushoff, J., Jennings, L., Smit, M., Macken, C., Hata, M., Gog, J., Miller, M.A., and Holmes, E.C. (2007). The genesis and spread of reassortment human influenza A/H3N2 viruses conferring adamantane resistance. *Mol Biol Evol* 24, 1811-1820.
- Simpson, P.A., Spudich, J.A., and Parham, P. (1984). Monoclonal antibodies prepared against Dictyostelium actin: characterization and interactions with actin. *J Cell Biol* 99, 287-295.
- Sontag, C.A., Sage, H., and Erickson, H.P. (2009). BtubA-BtubB heterodimer is an essential intermediate in protofilament assembly. *PLoS One* 4, e7253.
- Srinivasan, R., Mishra, M., Wu, L., Yin, Z., and Balasubramanian, M.K. (2008). The bacterial cell division protein FtsZ assembles into cytoplasmic rings in fission yeast. *Genes Dev* 22, 1741-1746.
- Stadler, J., Gerisch, G., Bauer, G., and Deppert, W. (1985). In vivo acylation of Dictyostelium actin with palmitic acid. *Embo J* 4, 1153-1156.
- Stein, L.D., Bao, Z., Blasiar, D., Blumenthal, T., Brent, M.R., Chen, N., Chinwalla, A., Clarke, L., Clee, C., Coghlan, A., Coulson, A., D'Eustachio, P., Fitch, D.H., Fulton, L.A., Fulton, R.E., Griffiths-Jones, S., Harris, T.W., Hillier, L.W., Kamath, R., Kuwabara, P.E., Mardis, E.R., Marra, M.A., Miner, T.L., Minx, P., Mullikin, J.C., Plumb, R.W., Rogers, J., Schein, J.E., Sohrmann, M., Spieth, J., Stajich, J.E., Wei, C., Willey, D., Wilson, R.K., Durbin, R., and Waterston, R.H. (2003). The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biol* 1, E45.
- Stoltzfus, A. (2006). Mutationism and the dual causation of evolutionary change. *Evol Dev* 8, 304-317.
- Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smailus, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A. (2002). Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences. *Proc Natl Acad Sci U S A* 99, 16899-16903.

- Stricker, J., and Erickson, H.P. (2003). In vivo characterization of *Escherichia coli* ftsZ mutants: effects on Z-ring structure and function. *J Bacteriol* **185**, 4796-4805.
- Suggs, S.V., Wallace, R.B., Hirose, T., Kawashima, E.H., and Itakura, K. (1981). Use of synthetic oligonucleotides as hybridization probes: isolation of cloned cDNA sequences for human beta 2-microglobulin. *Proc Natl Acad Sci U S A* **78**, 6613-6617.
- Takeda, M., Yoneya, A., Miyazaki, Y., Kondo, K., Makita, H., Kondoh, M., Suzuki, I., and Koizumi, J. (2008). *Prostheco bacter fluviatilis* sp. nov., which lacks the bacterial tubulin btubA and btubB genes. *Int J Syst Evol Microbiol* **58**, 1561-1565.
- Taoka, A., Asada, R., Wu, L.F., and Fukumori, Y. (2007). Polymerization of the actin-like protein MamK, which is associated with magnetosomes. *J Bacteriol* **189**, 8737-8740.
- Tebbich, S., Sterelny, K., and Teschke, I. (2010). The tale of the finch: adaptive radiation and behavioural flexibility. *Philos Trans R Soc Lond B Biol Sci* **365**, 1099-1109.
- Teng, G., and Papavasiliou, F.N. (2007). Immunoglobulin somatic hypermutation. *Annu Rev Genet* **41**, 107-120.
- Theissen, G. (2006). The proper place of hopeful monsters in evolutionary biology. *Theory Biosci* **124**, 349-369.
- Theissen, G. (2009). Saltational evolution: hopeful monsters are here to stay. *Theory Biosci* **128**, 43-51.
- Thomas, R.D., Shearman, R.M., and Stewart, G.W. (2000). Evolutionary exploitation of design options by the first animals with hard skeletons. *Science* **288**, 1239-1242.
- Thorburn, W.M. (1915). Occam's razor. *Mind*, 287—288.
- Toth-Petroczy, A., and Tawfik, D.S. (2011). Slow protein evolutionary rates are dictated by surface-core association. *Proc Natl Acad Sci U S A* **108**, 11151-11156.
- Towbin, H., Staehelin, T., and Gordon, J. (1979). Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. *Proc Natl Acad Sci U S A* **76**, 4350-4354.
- Travisano, M., and Velicer, G.J. (2004). Strategies of microbial cheater control. *Trends Microbiol* **12**, 72-78.
- Turner, N.J. (2003). Directed evolution of enzymes for applied biocatalysis. *Trends Biotechnol* **21**, 474-478.
- Udugama, M., Sabri, A., and Bartholomew, B. (2011). The INO80 ATP-dependent chromatin remodeling complex is a nucleosome spacing factor. *Mol Cell Biol* **31**, 662-673.
- Urban, M.C., Phillips, B.L., Skelly, D.K., and Shine, R. (2007). The cane toad's (*Chaunus* [Bufo] marinus) increasing ability to invade Australia is revealed by a dynamically updated range model. *Proc Biol Sci* **274**, 1413-1419.

- Uyeda, J.C., Hansen, T.F., Arnold, S.J., and Pienaar, J. (2011). The million-year wait for macroevolutionary bursts. *Proc Natl Acad Sci U S A* 108, 15908-15913.
- van den Ent, F., Moller-Jensen, J., Amos, L.A., Gerdes, K., and Lowe, J. (2002). F-actin-like filaments formed by plasmid segregation protein ParM. *Embo J* 21, 6935-6943.
- Vergara-Silva, F. (2003). Plants and the conceptual articulation of evolutionary developmental biology. *Biol. Philos.* 18, 249–284.
- Wada, A., Fukuda, M., Mishima, M., and Nishida, E. (1998). Nuclear export of actin: a novel mechanism regulating the subcellular localization of a major cytoskeletal protein. *Embo J* 17, 1635-1641.
- Wagner, G.P., Chiu, C.H., and Hansen, T.F. (1999). Is Hsp90 a regulator of evolvability? *J Exp Zool* 285, 116-118.
- Wang, H., Nussbaum-Wagler, T., Li, B., Zhao, Q., Vigouroux, Y., Faller, M., Bomblies, K., Lukens, L., and Doebley, J.F. (2005). The origin of the naked grains of maize. *Nature* 436, 714-719.
- Wang, J., Hou, L., Awrey, D., Loomis, W.F., Firtel, R.A., and Siu, C.H. (2000). The membrane glycoprotein gp150 is encoded by the lagC gene and mediates cell-cell adhesion by heterophilic binding during Dictyostelium development. *Dev Biol* 227, 734-745.
- Wang, R.L., Stec, A., Hey, J., Lukens, L., and Doebley, J. (1999). The limits of selection during maize domestication. *Nature* 398, 236-239.
- Ward, G.E., Brokaw, C.J., Garbers, D.L., and Vacquier, V.D. (1985). Chemotaxis of *Arbacia punctulata* spermatozoa to resact, a peptide from the egg jelly layer. *J Cell Biol* 101, 2324-2329.
- Weiser, A.A., Wittenbrink, N., Zhang, L., Schmelzer, A.I., Valai, A., and Or-Guil, M. (2011). Affinity maturation of B cells involves not only a few but a whole spectrum of relevant mutations. *Int Immunol* 23, 345-356.
- Williams, K.L., and Newell, P.C. (1976). A genetic study of aggregation in the cellular slime mould *Dictyostelium discoideum* using complementation analysis. *Genetics* 82, 287-307.
- Wilson, G.G. (1991). Organization of restriction-modification systems. *Nucleic Acids Res* 19, 2539-2566.
- Wilson, G.G., and Murray, N.E. (1991). Restriction and modification systems. *Annu Rev Genet* 25, 585-627.
- Worsdorfer, B., Woycechowsky, K.J., and Hilvert, D. (2011). Directed evolution of a protein container. *Science* 331, 589-592.
- Wright, S.I., Bi, I.V., Schroeder, S.G., Yamasaki, M., Doebley, J.F., McMullen, M.D., and Gaut, B.S. (2005). The effects of artificial selection on the maize genome. *Science* 308, 1310-1314.
- Yim, H.J., Hussain, M., Liu, Y., Wong, S.N., Fung, S.K., and Lok, A.S. (2006). Evolution of multi-drug resistant hepatitis B virus during sequential therapy. *Hepatology* 44, 703-712.

Yoshida, M., Inaba, K., and Morisawa, M. (1993). Sperm chemotaxis during the process of fertilization in the ascidians *Ciona savignyi* and *Ciona intestinalis*. *Dev Biol* 157, 497-506.

Yuan, R., Primakoff, P., and Myles, D.G. (1997). A role for the disintegrin domain of cyritestin, a sperm surface protein belonging to the ADAM family, in mouse sperm-egg plasma membrane adhesion and fusion. *J Cell Biol* 137, 105-112.

Zhang, P., Gu, Z., and Li, W.H. (2003). Different evolutionary patterns between young duplicate genes in the human genome. *Genome Biol* 4, R56.

List of figures

Figure 1. Organization of the microtubular and the actin system during cytokinesis.....	11
Figure 2. Actin Filaments.....	12
Figure 3. Microtubules.....	13
Figure 4. Intermediate Filaments.....	14
Figure 5. The impact of actin binding proteins on polymerization kinetics and stability of the microfilament network.....	15
Figure 6. Relatively frequent mutations in the actin molecule.....	18
Figure 7. Scanning electron micrograph showing the different developmental stages of <i>D.discoideum</i>	20
Figure 8. The corresponding locations of the actin genes on chromosomes 2 and 5.....	46
Figure 9. The evolutionary relatedness of the members of the actinome can be inferred from this tree.....	47
Figure 10. The alignment shows the insertions and deletions (indels) of Act8 as compared to the Arps.....	52
Figure 11. A tree analysis shows gross duplication events of the genes which code for identical actin proteins.....	53
Figure 12. Actin crystal structure, ATP-binding motifs, and modeled Act8 structure.....	55
Figure 13. Motif logos from <i>D. discoideum</i> actins and Arps.....	56
Figure 14. Motif logos from Act8 homologs across species.....	57
Figure 15. Structural homology among the <i>D. discoideum</i> Arps.....	59
Figure 16. A wider view of actin homologs from all available organisms.....	69
Figure 17. Distinct regions of Arp5 for an increased specificity of the antibodies generated....	73
Figure 18. The scheme of the construct of Arp5 fused to GST on a pGEX-6p1 vector.....	74
Figure 19. The expression pattern of the bacterial clones #1 and #10.....	75
Figure 20. Expression of fusion protein in clone #10 over 4h of induction.....	76
Figure 21. Test of the polyclonal antiserum.....	77
Figure 22. Detection of Arp5 with affinity purified polyclonal antibodies.....	78
Figure 23. Immunofluorescence with Arp 5 antibody on fixed cells.....	80
Figure 24. Schematic representation of the Arp5 knockout strategy.....	81
Figure 25. In this alignment the Arp6 region highlighted in yellow was used for proteinexpression.....	82
Figure 26. The scheme of the construct of Arp6 fragment fused to GST.....	83
Figure 27. The expression pattern of the Arp6 fragment fused to GST.....	84
Figure 28. Localization of GFP-Arp6 in AX2 wildtype cells.....	85
Figure 29. A pictorial description of the Arp6 knockout strategy.....	86
Figure 30: A depiction of micro- and macro-evolution in terms of fossil records.....	106

List of tables

Table 1. The major members of the actinome.....	43
Table 2. A list of actin pseudogenes.....	43
Table 3. The chromosomal locations of the identical actins.....	44
Table 4. The putative nuclear export signals.....	61
Table 5. The members of the actinome that harbor PEST motifs.....	64
Table 6. The <i>D. discoideum</i> actinome as compared to all currently available genomes.....	65
Table 7. The presence of the Arp members in different organisms.....	68
Table 8. ActA1 mutations which are found in members of the <i>D. discoideum</i> actinome.....	71

Abbreviations:

°C	Celcius
ABP	Actin Binding Protein
Act	Actin
AD	Adenosine binding loop
AID	Activation Induced Deaminase
Arp	Actin related protein
ATP	Adenosine Tri phosphate
BCIP	5-Bromo-4-chloro-3-indolylphosphate-p-toluidinsalz
BLAST	Basic Local Alignment Tool
Bp	Base pair
BSA	Bovine Serum Albumin
C	Connection motif
Ca	Calcium
cAMP	Cyclic Adenosine Mono Phosphate
cDNA	Complementary Deoxyribo Nucleic Acid
CIP	Calf Intestinal Phosphatase
Cm	Centimeter
DAPI	4, 6-Di-Amidino-2-Phenyl-Indole
dH ₂ O	Distilled water
DMSO	Di-Methyl Sulph Oxide
DNA	Deoxyribo Nucleic Acid
DNA-ML	DNA-Maximum Likelyhood
dNTP	Deoxy Nucleotide Tri Phosphate
DTT	1,4-Di-Thio-D,L-Threitol
E-value	Expectancy value

EB	Elusion Buffer
EDTA	Ethylene Diamine Tetra Acetic acid
EGTA	Ethylene Glycol bis [2-aminoethyl ether]--N,N,N',N'-tetraacetic acid
F-actin	Filamentous actin
G	Grams
G-actin	Globular actin
GC	Guanine Cytosine
GFP	Green Fluorescent Protein
GST	Glutathione S Transferase
GTP	Guanosine Tri Phosphate
H	Hours
HCl	Hydro Chloric acid
HSP	Heat Shock Protein
IF	Intermediate Filament
IgG	Immunoglobulin G
INO80	Inositol-requiring protein 80
IPTG	Iso-Propyl- β -D-Thio-Galactopyranoside
KDa	Kilo Dalton
kV	Kilo Volt
LB	Luria-Bertani media
M	Molar
MB	Mega Bases
MCS	Multiple Cloning Site
Mg	Magnesium
Min	Minutes
ml	Milli litre

Mm	Millimeter
mM	Millimolar
Ms	Millisecond
Mya	Million years ago
NCBI	National Centre for Biotechnology Information
NCP	Nucleus/Chromatin Preparation buffer
NES	Nuclear Export Signal
NGK	Nuclear Gate Keeper
NLS	Nuclear Localization Signal
Nm	Nano meter
OD	Optical Density
ON	Over Night
PAGE	Poly Acralamide Gell Electrophoresis
PBS	Phosphate Buffer Saline
PCR	Polymerase Chain Reaction
PEST	Proline (P), glutamic acid (E), serine (S), & threonine (T)
Pfam	Protein family database
PH	Phosphate binding motif
PIPES	PIPerazine-N,N'-bis(2-EthaneSulfonic acid)
PMSF	Phenyl Methyl Sulfonyl Fluoride
Rpm	Rotations per minute
S	Seconds
SBF	Species Branding Factors
SDS	Sodium Dodecyl Sulphate
SignalP	Signal Peptide detection server
SNP	Single Nucleotide Polymorphism

STET	Sucrose Triton X-100 EDTA Tris-HCl
Taq	Thermus aquaticus polymerase
TE	Tris EDTA
TM-HMM	Trans Membrane Hidden Markov Model
U	Units of enzyme
UV	Ultra Violet
V	Volume OR Volts
w/v	Weight per volume
μ F	Micro Faraday
ml	Micro Litre

Acknowledgements:

I am very grateful to Prof. Michael Schleicher for his constant support and guidance; especially more so for his kindness. I am also very fortunate to have had him as my doctor father, who (according to me) belongs to a rare breed of humans who are capable of thinking both through their brains and their hearts. But I find some of his characters to be infectious, for example his 'good sense of humor'; which successfully he had transmitted to me during my time in his lab!

I am very thankful to my friend Dr. Arasada Rajesh through whom I came to know about Prof. Schleicher and was able to join the lab. I am also thankful to my colleagues; Marlis Furbinger, Daniela Rieger, Dr. Nagendran Ramalingam, Dr. Petros Batsios, Dr. Annette Muller-Taubenberger, Dr. Peter Kastner and many others. I am also obliged to thank my friends in the Munich area namely, Sumathi, Sinnaiya Thavarasa, Selvakirubai, Mohan and Dr. Pandey and other good hearts.

I am also thankful to my earlier teachers and professors who have developed in me a thirst for life-sciences; namely Mr. Saravana Rajan, Mr. Radha Krishnan, Prof. S. Krishna Swamy, Prof. K. Dharmalingam. I also acknowledge the support I had received from all my friends around the world, many of whom, I unfortunately am not in constant touch. I am thankful to my friends Kannan Sriram, Yathesh Kumar, Mr. Gnanashekar and late Mr. Jawaharlal Paul.

I am greatly thankful to the Deutsche Forschungsgemeinschaft for funding my doctoral studies. Finally I am also thankful to all those who brought trouble upon me; without whom I would not have learned to steer away from paths which would have led me to troubles and dead ends.

Most of all, I cherish the support and motivation I received from my family members; my mother I. Daisy Mary, my father V.M Joseph, my sister Sophia Jegan, my uncle Prof. Henry Louis and other dear ones.

